Helsinki University of Technology Laboratory of Space Technology Publications
Teknillisen korkeakoulun Avaruustekniikan laboratorion julkaisuja
Espoo 2006 REPORT 67

Remote sensing of water quality for Finnish lakes and coastal areas

Sampsa Koponen

Dissertation for the degree of Doctor of Science in Technology to be presented with due permission of the Department of Electrical and Communications Engineering, for public examination and debate in Auditorium S4 at Helsinki University of Technology (Espoo, Finland) on the 22nd of December, 2006, at 12 noon.

Helsinki University of Technology
Department of Electrical and Communications Engineering
Laboratory of Space Technology

Teknillinen korkeakoulu
Sähkö- ja tietoliikennetekniikan osasto
Avaruustekniikan laboratorio

Distribution:
Helsinki University of Technology
Laboratory of Space Technology
P.O.Box 3000
FIN-02015 HUT
Finland

Tel. +358 9 451 2378
Fax +358 9 451 2898

Email: sampsa.koponen@tkk.fi
http://www.space.tkk.fi

ISBN-13 978-951-22-8533-4
ISBN-10 951-22-8533-9
ISBN-13 978-951-22-8534-1 (pdf)
ISBN-10 951-22-8534-7 (pdf)
ISSN 0786-8154

Picaset Oy
Helsinki 2006

Table of Contents

Preface

The research presented in this thesis was carried out in the Laboratory of Space Technology, Helsinki University of Technology (TKK), during 1997-2006. I owe my deepest gratitude to Professors Jouni Pulliainen and Martti Hallikainen for making my work possible and for advising me in my work. I would also like to thank the whole personnel of the laboratory for creating a wonderful working environment.

I am grateful to Kari Kallio, Jenni Attila (née Vepsäläinen), Timo Pyhälahti, Pekka Härmä, Yrjö Sucksdorff and other people at the Finnish Environment Institute (SYKE). Without your help this thesis would not have been done.

For financial support I would like to thank the Academy of Finland, TEKES, Jenny and Antti Wihuri foundation, Tekniikan edistämissäätiö and the GETA graduate school.

In the non-scientific community I would like to thank the "Krew" for all those good times we have had, "Thaikkuremmi" for many joyful occasions and all my friends around the world for just being there.

I thank my parents, sister and other relatives for the support and love they have given me over the years.

And finally Jonna, Luka and my new baby girl: Compared to you this thesis is insignificant.

Sampsa Koponen 1.12.2006

Abstract

In this thesis empirical remote sensing methods for estimating water quality in Finnish lakes and coastal areas are developed and tested. The remote sensing instruments used here are Airborne Imaging Spectrometer for Applications (AISA), Medium Resolution Imaging Spectrometer (MERIS) onboard the Envisat-satellite and Moderate Resolution Imaging Spectroradiometer (MODIS) onboard the TERRA-satellite.

Based on the results from this study the AISA airborne spectrometer is applicable for estimating chlorophyll a (chl *a*), Secchi depth, and turbidity in lakes. The 250-m resolution MODIS data are used for estimating turbidity and quality class in lakes. Full resolution (300 m) MERIS data are used for estimating chl a, total suspended solids (TSS) and the absorption coefficient of coloured dissolved organic matter (aCDOM(400)) during a spring bloom event in the Gulf of Finland (a situation where the default MERIS processor fails to provide valid data).

The retrieval of water quality information is based on single channel and channel-ratio algorithms, which are calibrated and tested with *in situ* (ground truth) observations. The accuracy of the retrieval is good. The results are based on a large number of data points (several thousand in one of the cases). Thus, the reliability of the results is high.

The thematic maps and statistics derived with remote sensing data demonstrate the advantages of remote sensing over the traditional water quality monitoring, which is based on *in situ* measurements. The main shortcoming of presented methods is that since the algorithms are based on empirical relationships, which include atmospheric effects, they require calibration (for different atmospheric parameters) before they can be used with other remotely sensed images. The effects of the atmosphere on MERIS channel-ratio algorithms are estimated with an atmospheric model.

Keywords: Remote sensing, water quality, MERIS, MODIS, AISA.

Tiivistelmä

Tässä väitöskirjassa on kehitetty ja testattu kaukokartoitusmenetelmiä, joilla voidaan arvioida Suomen järvien ja rannikkoalueiden vedenlaatua. Työssä käytetyt instrumentit ovat Airborne Imaging Spectrometer for Applications (AISA), Envisat-satelliitissa oleva Medium Resolution Imaging Spectrometer (MERIS) ja TERRA-satelliitissa oleva Moderate Resolution Imaging Spectroradiometer (MODIS).

Tulosten perusteella lentokonekäyttöisellä AISA spetrometrillä voidaan mitata klorofylli a:ta, näkösyvyyttä ja sameutta järvissä. 250-m erotuskyvyn MODIS-dataa voidaan käyttää järvien sameuden ja laatuluokan arviointiin. 300-m erotuskyvyn MERIS-dataa on käytetty klorofylli a:n, kiintoaineen sekä humuksen arvioitiin kevätkukinnon aikana Suomenlahdella (tilanne jossa Euroopan avaruusjärjestön käytössä oleva MERIS-prosessori ei pysty antamaan luotettavia tuloksia).

Vedenlaadun arvioiminen kaukokartoitusdatasta perustuu yksittäisten kanavien ja kanavasuhdealgoritmien käyttöön. Algoritmit on kalibroitu ja testattu maastohavaintoja käyttäen ja saatu mittaustarkkuus on hyvä. Tulokset perustuvat suureen datapisteiden määrään (useita tuhansia datapisteitä yhdessä tapauksessa), joten tulosten luotettavuus on korkea.

Kaukokartoitusdatan avulla aikaansaadut teemakartat ja tilastolliset tiedot osoittavat kaukokartoituksen edut tavanomaiseen vedenlaadun seurantaan verrattuna, joka perustuu maastohavaintoihin ja laboratoriomittauksiin. Työssä esitettyjen menetelmien suurin puute on se, että ne perustuvat empiirisiin yhtälöihin, jotka sisältävät ilmakehän vaikutukset signaaliin. Tämän vuoksi yhtälöt pitää kalibroida uudelleen, ennen kuin ne soveltuvat käytettäväksi muina ajankohtina otettujen kaukokartoituskuvien kanssa. Erilaisten ilmakehätilanteiden vaikutusta MERIS-kanavasuhdealgoritmeihin on arvioitu ilmakehämallin avulla.

Avainsanat: Kaukokartoitus, vedenlaatu, MERIS, MODIS, AISA.

List of Acronyms

6S	Second Simulation of the Satellite Signal in the Solar Spectrum
AISA	Airborne Imaging Spectrometer for Applications
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
AVHRR	Advanced Very High Resolution Radiometer
BOR	Bio-optical reflectance model
CASI	Compact Airborne Spectrographic Imager
CCD	Charge Coupled Device
CDOM	Colored dissolved organic matter
Chl *a*	Chlorohlyll a
CZCS	Coastal Zone Color Scanner
ESA	European Space Agency
ETM	Enhanced Thematic Mapper
EU	European Union
FMI	Finnish Meteorological Institute
FNU	Formazin Nephelometric Unit
HIRLAM	HIgh Resolution Limited Area Model
IR	Infrared
IRS	Indian Remote Sensing Satellite
LISS-III	Linear Imaging Self Scanning System
LST	Laboratory of Space Technology of the Helsinki University of Technology
MERIS	Medium Resolution Imaging Spectrometer
MODIS	Moderate Resolution Imaging Spectroradiometer
MODTRAN	MODerate spectral resolution atmospheric TRANSmittance algorithm and computer model, or Moderate Resolution Transmittance Code
NASA	National Aeronautics and Space Administration
NIR	Near Infrared
NOAA	National Oceanic and Atmospheric Administration
OC2, OC4	Ocean chlorophyll algorithms
OSS	Optically Significant Substance
RMSE	Root Mean Squared Error
SALMON	Satellite Remote Sensing for Lake Monitoring
SeaWiFS	Sea-viewing Wide Field Sensor
SPM	Suspended Particulate Matter
SPOT	Systeme Pour l'Observation de la Terre
SM	Suspended matter
SYKE	Finnish Environment Institute
TOA	Top of atmosphere
TM	Thematic Mapper
TSS	Total Suspended Solids

List of Symbols

a, b and c	Empirically determined regression parameters
$a(\lambda)$	Absorption coefficient
$a_{CDOM}(400)$	Absorption coefficient of water at 400 nm after filtering with a GF/F filter
$a_{ph}(\lambda)$	Absorption coefficient of phytoplankton
$a_s(\lambda)$	Absorption coefficient of suspended inorganic matter
$a_{tot}(\lambda)$	Total absorption coefficient
$a_w(\lambda)$	Absorption coefficient of water
$a_y(\lambda)$	Absorption coefficient of CDOM
$b(\lambda)$	Scattering coefficient
$b_b(\lambda)$	Backscattering coefficient
$b_{b,tot}(\lambda)$	Total backscattering coefficient
$b_{ph}(\lambda)$	Backscattering coefficient of phytoplankton
b_s	Backscattering coefficient of suspended inorganic matter
$b_w(\lambda)$	Backscattering coefficient of water
C	*C*oncentration of a water quality variable
C_{ph}	Concentration of phytoplankton
C_s	Concentration of suspended inorganic matter
C_y,	Concentration of CDOM
$E(\lambda)$	Spectral irradiance
$E_d(\lambda)$	Downwellin irradiance
E_{sun}	Solar Irradiance at the surface of Earth
$E_u(\lambda)$	Upwelling irradiance
f	A constant for environmental parameters
$L(\theta_i, \theta_r, \phi, \lambda)$	Radiance
L_a	Radiance from atmosphere
L_b	Radiance from bottom
$L_{ch\text{-}1}$	MODIS channel-1 (620-670 nm) radiance
L_d	Radiance detected at sensor
L_i,	Incident radiance
L_l	Radiance from adjacent areas
L_r	Reflected radiance
L_s	Radiance from water surface
L_t	Transmitted radiance

L_u	Upwelling radiance
L_w	Radiance from water
$L_{w,0}$	Radiance just above the water surface
$L_{w,-0}$	Radiance just below water surface
N	Number of data points
$R(\lambda)$	Irradiance reflectance
$R(0-,\lambda)$	Reflectance just below the surface of water
R_1 *and* R_2	Remotely sensed reflectances or radiances
R^2	Coefficient of Determination
R_{RS}	Radiance reflectance (or remote-sensing reflectance)
R_s	Reflectance of the water surface
SE_M	Standard error of a mean value of a remotely estimated parameter
T	Turbidity
T_{atm}	Atmospheric transmittance
T_s	Transmittance of water surface
ϕ	Azimuth angle between the incident and reflected radiation
λ	Wavelength
μ_0	Cosine of solar zenith angle under water surface
θ_i	Zenith angle for incident radiation
θ_r	Zenith angle for reflected radiation
σ	Standard deviation

List of Appended Papers

This thesis is based on the work contained in the following publications, hereafter referred to as publications [P1] to [P6]:

[P1] **Koponen, S.**, Pulliainen, J., Servomaa, H., Zhang, Y., Hallikainen, M., Kallio, K., Eloheimo, K., Pyhälahti, T., Hannonen, T., Analysis on the feasibility of multi-source remote sensing observations for chl *a* monitoring. The Science of the Total Environment, vol. 268, nos. 1-3, pp. 95-106, 2001.

[P2] Kallio, K., **Koponen, S.**, and Pulliainen, J., Feasibility of airborne imaging spectrometry for lake monitoring—a case study of spatial chlorophyll a distribution in two meso-eutrophic lakes. International Journal of Remote Sensing, Vol. 24, No. 19, pp. 3771–3790, 2003.

[P3] **Koponen, S.,** Attila, J., Pulliainen, J., Kallio, K., Lindfors, A., Rasmus, K., and Hallikainen, M., A Case Study of Airborne and Satellite Remote Sensing of a Spring Bloom Event in the Gulf of Finland. Continental Shelf Research, in press, 2006.

[P4] **Koponen, S.**, Pulliainen, J., Kallio, K., Vepsäläinen, J. and Hallikainen, M., Use of MODIS data for monitoring turbidity in Finnish Lakes. IEEE 2001 International Geoscience and Remote Sensing Symposium (IGARSS'01), CD-ROM, 3 p., Sydney, Australia, 9-13 July, 2001.

[P5] **Koponen, S.,** Pulliainen, J., Kallio, K., and Hallikainen, M., Lake water quality classification with airborne hyperspectral spectrometer and simulated MERIS-data. Remote Sensing of Environment, Vol. 79/1, pp. 51-59, 2002.

[P6] **Koponen, S.**, Pulliainen, J., Kallio, K., Vepsäläinen, J., Pyhälahti, T., and Hallikainen, M., Water quality classification of lakes using 250-m MODIS data, Geoscience and Remote Sensing Letters, vol. 1, no. 4, pp. 287- 291, 2004.

In [P1] and [P3]-[P6] the first author was responsible for carrying out the presented research. The other authors acted as scientific advisors, gave ideas and processed or produced data. In [P2] the research was performed in collaboration with K. Kallio. The bio-optical model used in [P6] was developed by K. Kallio. The atmospheric model used in [P3] and [P6] was developed by J. Attila (née Vepsäläinen).

1 Introduction

Water is a necessity for all living organisms. While some basic life forms thrive even in polluted water, most organisms require access to water that is relatively clean (i.e. has good quality). The quality of surface waters is affected by the type and the quantity of various suspended and dissolved substances (see Table 1.1). Generally, a large amount of these substances means poor water quality. Some of the substances occur naturally while others originate from human activities. Loading from such sources as agriculture, fish farming, industry and municipalities can cause eutrophication and have other adverse effects on water quality. For more information on the *in situ* (in the field) water quality measurements see Chapter 1.2.2.

Remote sensing can be defined as the science of using an instrument for measuring a target and its properties without a physical connection between the instrument and the target. Typically, the measurements are performed by using electromagnetic radiation (e.g. ultra-violet, visible light, reflective and thermal infra-red, and microwaves). The instrument records the radiation reflected or emitted by the target and its properties are then inferred from the measured signal.

One of the advantages of remote sensing is that the measurements can be performed from a great distance (several hundred or even several thousand km in the case of satellite sensors), which means that large areas on ground can be covered easily. With satellite instruments it is also possible to observe a target repeatedly; in some cases every day or even several times per day.

The aim of this thesis is to examine the use of remote sensing for determining water quality in Finnish lakes and coastal areas. The structure of this thesis is the following: The first chapter starts with a description of Finnish lakes and coastal areas. In the second part of Chapter 1 the monitoring of water quality is discussed. In Chapter 2 the theoretical basis of remote sensing of water quality and the factors that affect it are explained. In Chapter 3 the current state-of-the-art is reviewed. In Chapter 4 the objectives of this work are listed, the data used in the study are presented and the methods used to process and analyze the data are explained. The results obtained by the author are also presented in Chapter 4. In Chapter 5 their meaning is discussed. Finally, in Chapter 6 the matters that will be researched in the future are addressed.

Table 1.1. The parameters affecting water quality. See Chapters 2.3 and 2.4 for more details.

Parameters that can be estimated with optical remote sensing methods
In situ data analyzed in laboratory from water samples:
• Phytoplankton (chlorophyll *a*)
• Suspended inorganic material (e.g. sand, dust and clay)
• Colored dissolved organic matter
• Turbidity
• Biological status related to phytoplankton and other aquatic flora [1]
In situ data estimated in the field:
• Secchi depth
• Temperature
• Occurrence and extent of algal blooms
Parameters that cannot be estimated directly[2] with current optical remote sensing methods
• Toxic heavy metals (e.g. Hg, As, Cr, Pb and Cd)
• Toxic chemicals (e.g. dioxin)
• Algae species composition
• Bacteria (especially hygienic indicators such as fecal coliforms)
• Taste deficits of fish
• Nutrients (nitrogen and phosphorus)
• Taste and odour
• pH
• Dissolved oxygen content (top and bottom water layers)
• Chemical oxygen demand
• Biochemical oxygen demand
• Salinity and conductivity [3]
• Biological status related to aquatic fauna [1]

[1] See Chapters 1.2.2 and 1.2.3 for more details.

[2] Some indirect methods (e.g. using another parameter as an indicator of the desired one) may work in some cases.

[3] Microwave radiometers operating in the L-band (1.4 GHz) can be used to monitor sea surface salinity.

1.1 Natural waters in Finland

1.1.1 Lakes

Most Finnish lakes are from glacial origin; i.e. they were formed when the glacier of the latest ice age gradually melted about 10 000 years ago (Särkkä, 1996). The main characteristics of Finnish lakes are presented in Table 1.2.

During most winter the lakes in Finland have ice cover. This causes the biology of lakes and the evolution of water quality parameters to have a roughly repeating yearly cycle. During winter only a very small amount of radiation reaches the phytoplankton cells, which are the basic productive element in water. Therefore, the productivity of lakes is low or nonexistent. When the radiation levels rise as the ice cover melts in the spring the biological activity starts to increase. The nutrients released from decayed organisms and the bottom during winter and the increasing temperature support the growth. The amount of phytoplankton has a maximum some time after the ice cover has completely melted. When most of the available nutrients have been used and the amount of animal plankton (which graze on phytoplankton) grows the amount of phytoplankton starts to decrease. A minimum amount is reached during the summer. During autumn the nutrients stored in deeper water become available due to thermal mixing of water and another growth period is observed. The phytoplankton species responsible for the growth peaks observed during spring and autumn can be different. The decreasing amount of animal plankton (due to decreasing temperature) also contributes to the growth. Then the water surface freezes and the cycle starts again. Figure 1.1 shows the cycle in graphic form.

In lakes with low nutrient content only one peak is observed and the maximum amount of phytoplankton occurs during summer. In some lakes the behavior of other water quality variables such as Secchi depth and turbidity (see chapter 2.4 for definitions) follows the cycle of phytoplankton while in other it does not.

Finnish lakes can be divided into trophic classes according to e.g. the amount of nutrients available for plant growth. Table 1.3 gives an indication of the values of the class limits and the estimated portion of lakes belonging to each class. Lakes can also be classified according to the amount of humic compounds they contain. Finnish lakes are very humic when compared to lakes in rest of the world.

Table 1.4 shows the limits of some water quality variables in Finnish lakes using routinely collected data. As can be seen the variation range of the values is substantial. Some of this variation is lake dependent and some is due to seasonal changes.

Table 1.2. Characteristics of Finnish lakes (Raatikainen and Kuusisto, 1988), (Ympäristö 2004), (Ympäristö 2006).

Number of lakes (larger than 0.01 km^2)	56 012
Total area covered by lakes	33 350 km^2 (approximately 10 % of Finland)
Total volume (larger than 0.01 km^2)	235 km^3
Estimated mean depth	7 m
Maximum depth	96 m (in Lake Päijänne)

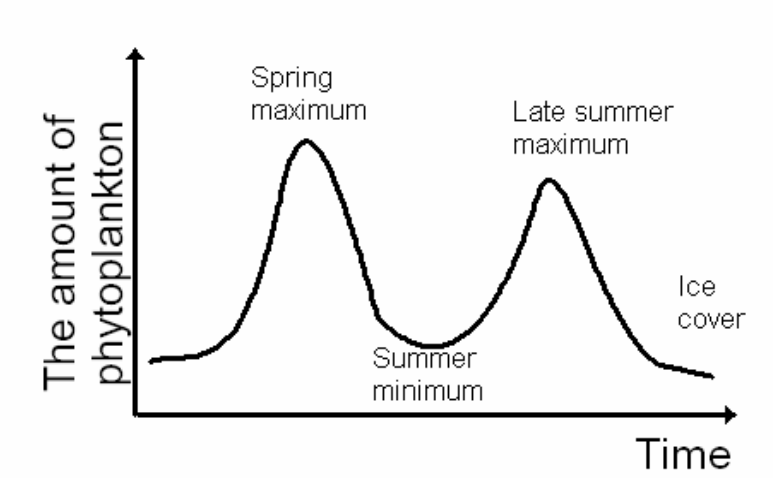


Figure 1.1. The yearly cycle of the variation of phytoplankton as a function of time (the magnitudes and the locations of the peaks depend on the trophic class of the water body and other factors).

Table 1.3. The classification of Finnish lakes into trophic classes. The values presented here are not absolute; they depend on such factors as the size and the shape of the lake basin and the variation within the yearly cycle (Särkkä, 1996).

	Total phosphorous (μg/l)	Chl *a* (μg/l)	Basic production (mg C/m^2/day)	Estimated portion of lakes* (%)
Ultraoligotrophic		<1	<50	
Oligotrophic	<15	1-5	10-100	38.5
Mesotrophic	15-25	5-10	100-250	52.5
Eutrophic	>25	10-50	250-1000	9.0
Hypereutrophic		>1000		

* From Henriksen et al. (1997); Total phosphorous limits: oligotrophic <10 μg/l and eutrophic >35 μg/l.

Table 1.5 shows the correlation coefficients between various water quality parameters for lake data. All parameters except Secchi depth are positively correlated with each other. TSS is highly correlated with chl *a*, turbidity and inverse Secchi depth. The correlation between $a_{CDOM}(400)$ and the other parameters is low. The correlation between TSS and chl a, and TSS and Secchi depth are shown in graphic form in Figure 1.2 and Figure 1.3.

There are over 500 000 lakes in Europe larger than 0.01 km^2 (EEA, 2006) and almost half of the large (> 100 km^2) European lakes are located in Finland. When compared to other European lakes, the ones most similar to Finnish lakes can be found in Sweden as they were also formed after the latest ice age. The Swedish classification limits for Secchi depth, total phosphorous and chl *a* (chlorophyll *a*) are similar to those used for Finnish lakes (Swedish EPA, 2006). On the other hand, for turbidity the Swedish limits are lower than the Finnish limits.

*Table 1.4. The minimum, maximum, mean and median values of water quality parameters at lake and coastal routine monitoring (see chapter 4.2) stations during 2000-2005. N is the number of data points for each parameter. Lines marked with * are from samples collected from 1 m depth. With the rest composite samples are used. In lakes the composite samples are collected from 0 to 2 m depth and mixed before analysis. In coastal stations the depth interval starts from 0 m and the end value depends on the visibility (sampling depth can be up to 16 m). The number of stations is 5709 and 1310 for lakes and coast, respectively. TSS is Total suspended solids, $a_{CDOM}(400)$ is the absorption coefficient of CDOM at 400 nm.*

Lakes	Min	Max	Mean	Median	N
Chl *a* (μg/l)	0.2	730	15.3	8.5	23308
Turbidity (FNU)*	0.03	220	3.8	1.9	22642
Secchi Depth (m)	0.05	29	2.0	1.8	26530
TSS (mg/l)	0.0	150	2.6	1.3	1151
$a_{CDOM}(400)$	0.092	35.9	6.0	4.6	1034
Coast					
Chl *a* (μg/l)	0.05	500	8.5	5.2	13746
Turbidity (FNU)*	0.05	630	4.56	1.9	9926
Secchi Depth (m)	0.03	28	2.4	2.2	16081
TSS (mg/l)*	0.2	39	3.3	2.1	1331
$a_{CDOM}(400)$*	0.013	28.6	2.1	1.5	1161

Table 1.5. Correlation coefficients (r) between water quality parameters. For Secchi depth the correlation coefficient is also given for the inverse relationship (e.g. TSS vs. 1/Secchi) in parentheses. The values are computed from data collected from routine monitoring stations during 2000-2005 (see chapter 4.2 and Table 1.4)

	Turbidity (FNU)	Secchi Depth (m)	TSS (mg/l)	$a_{CDOM}(400)$
Chl *a* (μg/l)	0.57	-0.41 (0.59)	0.83	0.39
Turbidity (FNU)		-0.36 (0.72)	0.86	0.27
Secchi Depth (m)			-0.54 (0.87)	-0.61 (0.52)
TSS (mg/l)				0.28

A comparison of water quality in European lakes was presented by Kristensen and Hansen (1994). They used the concentration of total phosphorous as an indicator of water quality in different parts of Europe and noted that in Finland and Sweden the concentration is relatively low (below 25 μg/l for 75-90 % of the lakes). In Norway the concentration is less than 10 μg/l for 70 % of the lakes. In Germany and Italy the concentration is between 25 and 125 μg/l for about 50 % of the lakes. Alpine lakes generally have a lower phosphorous concentration. In England the concentration is over 50 μg/l in 80 % of the lakes. These numbers, although they do not represent all important water quality variables, give an indication of the variability of water quality variables in Europe.

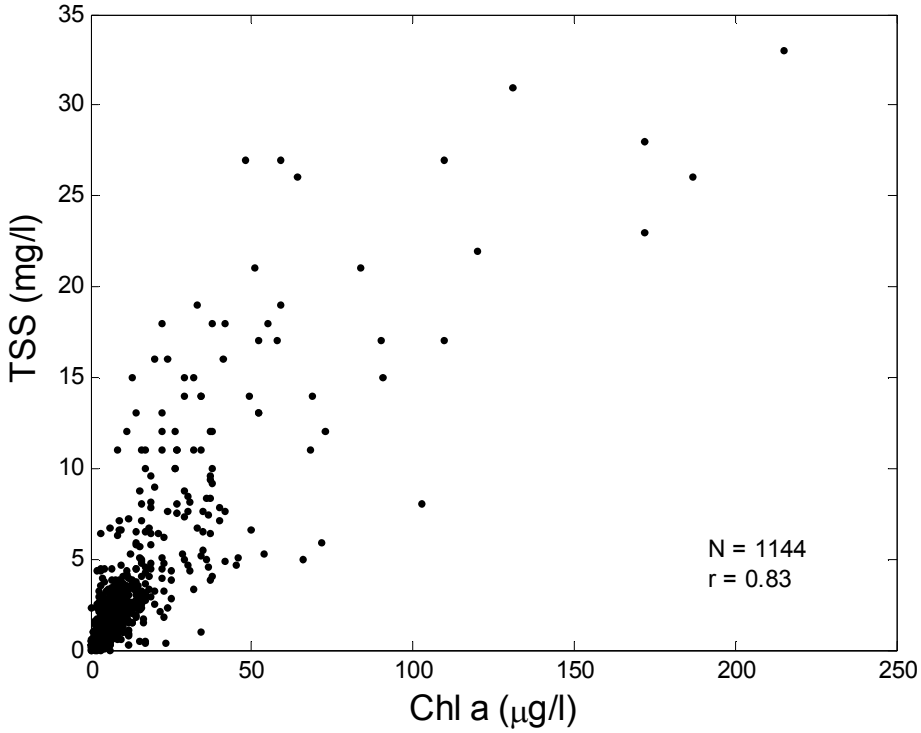


Figure 1.2. Chl a vs. TSS using routinely collected lake data from Finland (years 2000-2005). N is the number of data points. r is the correlation coefficient.

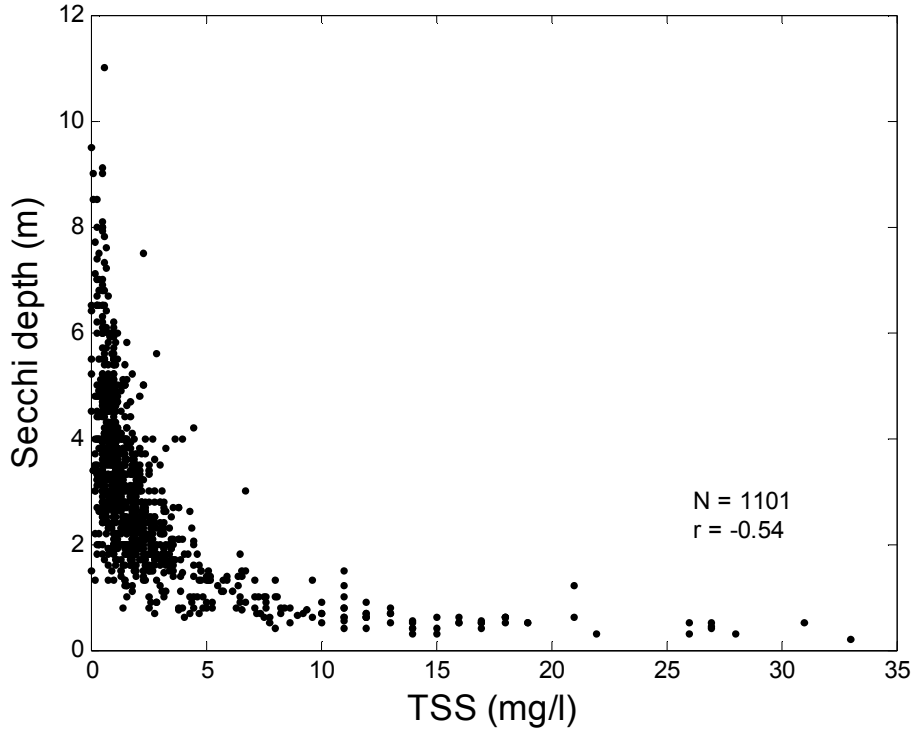


Figure 1.3. TSS vs. Secchi depth using routinely collected lake data from Finland (years 2000-2005). N is the number of data points. r is the correlation coefficient.

1.1.2 Coastal areas

Finland has a coast with only one sea, the Baltic Sea (map shown in Figure 1.4). Table 1.6 shows the characteristics of Finnish coastal areas. The water in the Baltic Sea is brackish, i.e. its salinity varies. In fact, Baltic Sea is the largest pool of brackish water in the world. Baltic Sea is quite shallow and due to its small volume Baltic Sea is susceptible to pollution (mainly agricultural and municipal loading). Cyanobacterial blooms have become common in some areas during late July and August.

The easternmost part of the Baltic Sea is called Gulf of Finland. Gulf of Finland is an active shipping route and due to its low depth it is even more susceptible to pollution than the rest of the Baltic Sea.

The quality of coastal water can have large spatial and temporal variations. Similarly to lakes, the coastal areas of Finland have ice cover during most winters. The southern parts of the Baltic Sea, on the other hand, do not usually freeze. The spring and late summer blooms also occur regularly, just like in lakes. Rivers can have a large effect on water quality in a local scale, especially during the snow melting season, as the melt water brings nutrients and inorganic matter to the coast.

The minimum and maximum values of some water quality parameters measured at stations along the coast of Finland are shown in Table 1.4. As with lakes, the variability range is large.

Table 1.6. Characteristics of Finnish coast (Perttilä et al. 2004), (Ahlman 2003), (Ympäristöministeriö 2006).

Direct length	1 100 km
Length including islands, bays and peninsulas	46 000 km
Surface area of Baltic Sea	422 000 km^2
Surface area of Gulf of Finland	29 570 km^2
Volume of Baltic Sea	21 000 km^3
Volume of Gulf of Finland	1 103 km^3
Average depth of Baltic Sea	55 m
Maximum depth of Baltic Sea	450 m
Average depth of Gulf of Finland	37 m
Maximum depth of Gulf of Finland	123 m

Figure 1.4. Map of the Baltic Sea. Lakes are also shown. The map is drawn with a service developed by UNEP/GRID-Arendal (http://maps.grida.no/baltic/).

1.2 Water quality monitoring

1.2.1 Reasons for water monitoring

The monitoring of the quality of water in lakes, coastal areas and rivers is required by national laws[1] and directives[2,3] of the European Union (EU). The main objective of these regulations is to prevent and control the pollution of the environment.

[1] Environmental protection law 86/2000

[2] Integrated Pollution Prevention and Control (IPPC) Directive, Council Directive 96/61/EC (September 24, 1996)

[3] EU Water Framework Directive 2000/60/EC

Environmental monitoring is one of the methods that contribute to reaching this goal. For example, water management officials and planners need up to date information on the current status of the watercourses and the effect of human activity on them so that they can identify pollution sources and act accordingly. Lakes are a major source of raw water for domestic and industrial use in Finland and thus monitoring lake water quality is important for economical reasons, as clean water does not need to be purified as much as water polluted with, e.g., blue-green algae. Lakes and coastal areas are also important places for other economical activities such as fishing. Also the general public is interested in the state of water, especially near their summer cottages and public swimming places.

1.2.2 Traditional method

In Finland the quality of lake, river and coastal waters is traditionally determined by collecting water samples and analyzing the samples in a laboratory (e.g. chlorophyll *a*, turbidity and total phosphorous, see Chapter 4.2.1 for the details on the methods used in the laboratory), and by making on-site measurements (e.g. Secchi depth and temperature). However, the sampling is slow and expensive and it does not include all lakes. In fact, less than 3000 lake and less than 1000 coastal stations are sampled each year. The number of visits per station varies. There currently are 28 so called intensive monitoring stations that are visited roughly every two weeks (the ones in Lapland are only visited four times per year). The rest, a vast majority, are visited only once or a few times per year. Also, information on the spatial variances of water quality within lakes is limited. Given the number of the lakes the monitoring is a huge task.

The EU Water Framework Directive (WFD) 2000/60/EC requires changes in the way water areas are monitored. The emphasis will be on assessing the biological status of water instead of using purely physical and chemical parameters. The biological elements include:

- Composition, abundance and biomass of phytoplankton
- Composition and abundance of other aquatic flora
- Composition and abundance of benthic invertebrate fauna
- Composition, abundance and age structure of fish fauna

The monitoring of these elements is even more difficult, time consuming and expensive than the monitoring of physical and chemical parameters. Thus, the directive will increase the amount of resources needed for water monitoring.

1.2.3 Remote sensing

Some of the substances found in water contribute to the way in which optical radiation interacts with water bodies (through wavelength-dependent scattering and absorption, see Chapters 2.1 and 2.2 for more details). In a sense, these optically significant substances change the color of water. Remote sensing instruments using certain channels in the optical region of the spectrum can detect the changes and estimate the amount of optically significant substances (see Chapter 2 for more details).

The research on the remote sensing of water quality has concentrated on open seas (ocean color studies) leaving coastal and especially inland water areas with much less attention. The lack of suitable sensors has limited the use of remote sensing in the past especially for lakes.

The first satellite based sensor devoted to water quality measurements was the Coastal Zone Color Scanner (CZCS) (Hovis et al., 1980), launched in 1978. Since then the availability of remote sensing data suitable for monitoring water quality has improved, especially recently. SeaWiFS (Sea-viewing Wide Field Sensor, launched in 1997) continued on the path started by CZCS and has also produced good results (e.g. Moore et al., 1999). Airborne instruments such as AISA (Airborne Imaging Spectrometer for Application) (Mäkisara et al., 1993), CASI (Compact Airborne Spectrographic Imager) (Babey and Anger, 1989) and HyMap (Cocks et al., 1998) have been determined to be feasible for monitoring small areas in Scandinavia (Kallio et al., 2001, Östlund et al., 2001) and Germany (Thiemann and Kaufmann, 2000b). Data from airborne sensors can also be used for developing retrieval algorithms for spaceborne sensors such as MODIS (Barnes et al., 1998) and MERIS (Medium Resolution Imaging Spectrometer) (Rast et al., 1999).

Spaceborne sensors designed for other applications have also been useful for water quality monitoring. For example, Landsat TM data have been used for the estimation of total suspended solids (Dekker and Peters, 1993) and chlorophyll (Mayo et al., 1995). NOAA AVHRR data have been used for monitoring light attenuation (a measure of turbidity) in a coastal area by Woodruff et al. (1999).

Current remote sensing methods are able to monitor only some of the biological elements mentioned in Chapter 1.2.2. These are the composition (on a very coarse level), abundance and biomass of phytoplankton, and the composition (on a very coarse level) and abundance of other aquatic flora. The elements involving benthic invertebrate and fish fauna cannot be monitored at all with current remote sensing technology.

Even though remote sensing methods cannot monitor all of the required parameters or elements, they are able to monitor some of them with good spatial and temporal coverage. Thus, they can one day – once the methods are mature enough – provide more resources for water quality monitoring, which allows the water quality experts to concentrate on the biological side.

The methods used in the remote estimation of water quality can be divided into empirical and analytical (a combination is also possible). The empirical methods relate the remote sensing signal to the *in situ* data by some statistical method of minimizing the difference or error between the variables. The resulting algorithm is valid only for the remote sensing image that is used in the training of the algorithm. This usually restricts the use of empirical methods to those cases where *in situ* data are available. Analytical methods use bio-optical models, which are based on the basic interactions (scattering and absorption) between water and radiation and thus are generally more valid for different areas and water types. One problem with analytical methods is that if the parameter values of the model (see Chapter 2.3) are not correct, the estimation of the water quality parameters can have large errors.

1.2.4 Oceans vs. lakes and coastal areas

Natural waters can be divided into two classes (Morel and Prieur, 1977; Gordon and Morel, 1983; Morel, 1988)). The so-called Case 1 waters include the (usually mid-oceanic) areas, where the dominating optically significant substance is phytoplankton (other optically significant substances and their influence on the optical properties of water covary with phytoplankton). The Case 2 waters in turn include optically more complex water areas such as lakes and coastal areas where other optically significant substances such as suspended inorganic matter and colored dissolved organic matter (CDOM) can have a significant or even dominating effect on the optical properties of water independently from the concentration of phytoplankton. Although it has recently been proposed (Mobley et al. 2004) that the simple classification of water into Case 1 and Case 2 should be dropped, the classification is still employed here.

It is possible that the water in a large clear lake is optically similar to Case 1 water. However, for Finnish lakes this is rarely (if ever) true and it is safer to work under the Case 2 assumption.

The situation is the same for the Baltic Sea. It can be difficult to define where the coastal area ends and where open sea starts. For example, the WFD defines the coastal zone to be the area within one nautical mile from the outer islands of the archipelago. This artificial limit based on distance from the shore is not very good for the Baltic Sea, where cyanobacterial blooms often occur in the center areas far from land. Thus, in this work the definition of a coastal area is extended to include all areas where the Case 2 assumption can be assumed to be true. Hence, the whole Baltic Sea is considered to be coastal area.

The remote sensing algorithms developed for Case 1 waters (starting from the CZCS algorithms, Gordon and Clark, 1980) are widely accepted and are fairly reliable, although local modifications may be needed in some cases (Jorgensen 1999). Global water quality maps based on these algorithms are already available. Unfortunately, that is not the situation for Case 2 waters. Due to the complexity of the composition of Case 2 water it may be difficult to separate from each other the optical signals caused by different substances, which form the remotely detected signal. Also, while the development of remote sensing methods has been rapid the retrieval algorithms are not globally applicable. Methods that work well for one target area regularly fail for another. Therefore, area-specific monitoring systems are often required.

Atmospheric correction is also more complicated for Case 2 than it is for Case 1. Further, the shape of lakes is often irregular, especially for post glacial lakes such as Finnish lakes. They often have a complex system of peninsulas, bays, straits and islands, which also hamper the retrieval, particularly with instruments that have low spatial resolution. A buffer zone of one or more pixels is usually needed at the shore of a lake in order to remote pixels contaminated with land or shore vegetation. Hence, the distance between the opposite shores of a lake at its narrowest point has to be at least 3 to 5 times the pixel size of the instrument before the lake can be measured with it. For MERIS and MODIS type of instruments, this means that the lake has to be at least 1 km wide. For instruments similar to Landsat-TM the lake would have to be at least 100 m wide.

2 Theoretical background

In this chapter the basics of remote sensing of water are presented. The factors affecting the measurement are discussed and the theoretical basis for the optical nature of water is explained.

2.1 Basics of remote sensing

Remote sensing measurements are usually based on interactions between electromagnetic radiation and matter. In remote sensing of water quality the most important region of the electromagnetic spectrum is the visible band (roughly 400 to 700 nm). The attenuation of infrared (IR, roughly 700 nm to 15 μm) and microwave radiation (1 mm to 30 cm) is so high in water (Zoloratev and Demin, 1977) that those wavelengths can only be used for observing surface phenomena such as algae blooms.

The most common remote sensing instrument used for water monitoring is a passive spectrometer operating in the optical region of the spectrum. Passive instruments measure the solar radiation that is reflected by water as opposed to e.g. radar, which generates its own radiation. Spectrometers observe the reflected radiation in several wavelength channels (also known as bands) covering a certain portion of the visible-infrared region. Other instruments that can be used for water quality monitoring include microwave radars. However, those can only observe surface phenomena as discussed above. Lidars (laser radars) operating in the optical region can be used to retrieve water quality information from below the surface but the availability of lidar data is poor. Hence, these two instrument types will not be covered in this work

2.2 Remote sensing of water quality

Five radiation components may contribute to the total intensity measured by a passive sensor (see Figure 2.1). The first component originates from the sun and propagates through the atmosphere and through the air-water interface into water. In water, the radiation is scattered and absorbed by water molecules and substances that are suspended or dissolved in water. A portion of the radiation is scattered towards the surface. This upwelling radiation then travels through the air-water interface and at least a portion of the atmosphere (depending on the altitude of the sensor) and finally reaches the sensor. All these steps have an effect on the intensity and the spectral shape of the radiation and, therefore, affect the estimation of water quality.

The other four other radiation components (2 to 5) shown in Figure 2.1 contain no information about the water quality, and hence act as noise to the measurement hindering the estimation of water quality. As we shall later see, the intensity of the radiation reflected from the water volume is typically low and the other components can dominate the detected signal. These components are described in more detail in Chapters 2.2.1 - 2.2.4.

In the field of remote sensing the most commonly used measure of radiant energy is radiance (L). Radiance is defined as the amount of radiant energy per unit time per unit solid angle (i.e. towards a certain direction) per unit of projected area of the source (i.e. its unit is W m^{-2} sr^{-1}). Spectral radiance is the radiance per unit wavelength

interval at a given wavelength (unit W m^{-2} sr^{-1} μm^{-1}). The radiance detected by the sensor in Figure 2.1 (L_d) can be expressed as

$$L_d = T_{atm} (L_w + L_s + L_b) + L_a + L_l, \quad (2.1)$$

where L_w, L_s, L_a, L_b, and L_l are the radiance components from the water volume (at depth z = +0, just above the surface), water surface (depth z = +0), atmosphere, bottom and adjacent areas, respectively, and T_{atm} is the transmittance of the atmosphere (the portion of the radiation that propagates through the atmosphere). Each radiance component is a function of the measurement geometry (i.e. solar and sensor angles) and wavelength (λ) and can be expressed as $L(\theta_i, \theta_r, \phi, \lambda)$, where θ_i and θ_r are the zenith angles for incident and reflected radiation, respectively, and ϕ is the azimuth angle between the incident and reflected radiation (see Figure 2.2 for the definitions of these angles).

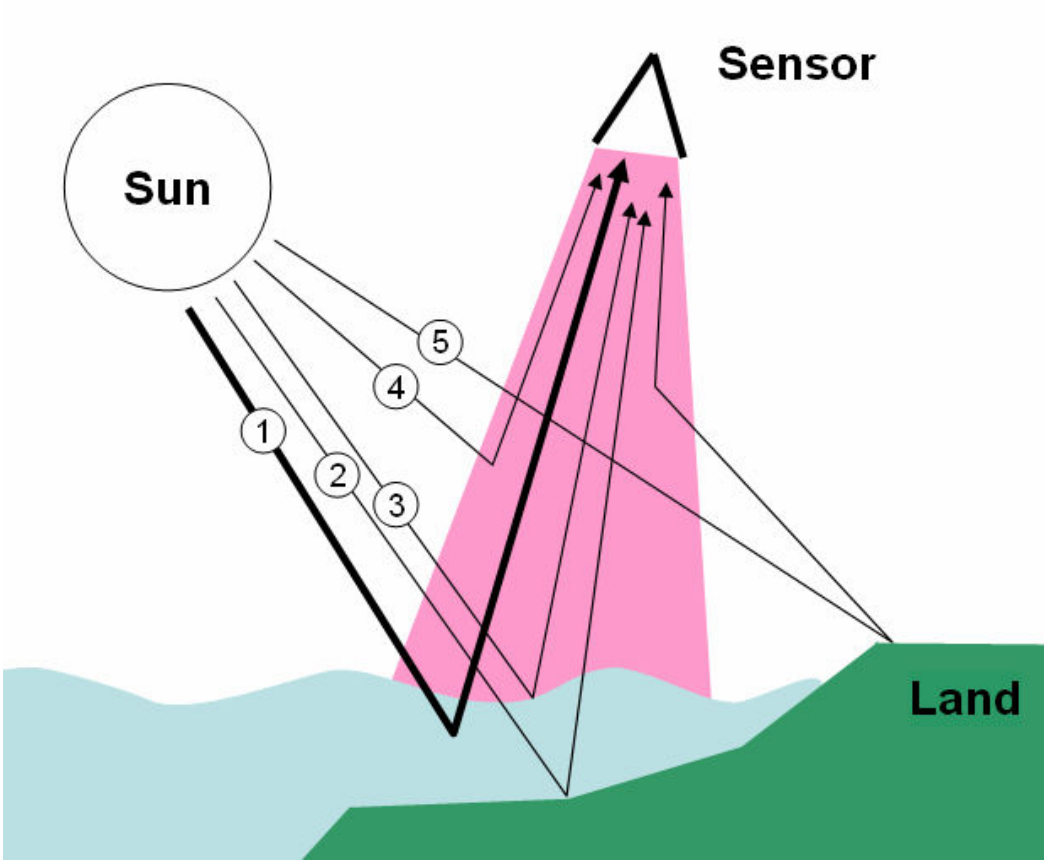


Figure 2.1. The components of radiance reaching a passive sensor (field-of-view indicated with pink color). (1) Radiation upwelling from the water volume. (2) Radiation reflected from the bottom. (3) Radiation reflected by the air-water interface. (4) Radiation scattered to sensor by the atmosphere between water and sensor. (5) Radiation reflected from a target that is close to the field-of-view of the sensor and scattered to the sensor by the atmosphere.

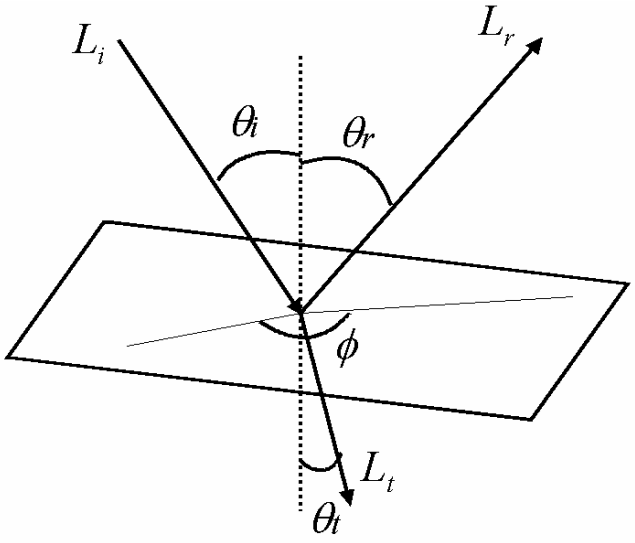


Figure 2.2. The geometry of a remote sensing measurement. L_i, L_r and L_t are the incident, reflected and transmitted radiances, respectively, θ_i and θ_r are the zenith angles for incident and reflected radiation, respectively, and ϕ is the azimuth angle between the incident and reflected radiance.

Radiance is closely related to irradiance $E(\lambda)$, which is defined as the flux of radiant energy per unit surface area (unit W m^{-2}). I.e., E includes radiation originating from all directions of a half-sphere. The portion of the irradiance that is heading to all downward directions is called the downwelling irradiance $E_d(\lambda)$. The upward heading portion of the irradiance in turn is called the upwelling irradiance $E_u(\lambda)$. Irradiance reflectance $R(\lambda)$ is defined as the ratio of upwelling and downwelling irradiances:

$$R(\lambda) = \frac{E_u(\lambda)}{E_d(\lambda)}. \tag{2.2}$$

Since remote sensing instruments measure radiance it is often more convenient to use radiance reflectance R_{RS} (also know as the remote-sensing reflectance), which can be expressed as

$$R_{RS}(\lambda) = \frac{L_u(\lambda)}{E_d(\lambda)}. \tag{2.3}$$

where L_u is the upwelling radiance component including the contributions from water volume, water surface and bottom ($L_u = L_w + L_s + L_b$). The main difference between $R(\lambda)$ and $R_{RS}(\lambda)$ is that for $R_{RS}(\lambda)$ the direction of the upwelling radiation (L_u) is towards a small solid angle while for $R(\lambda)$ a half-sphere is used.

2.2.1 Atmosphere

The most evident atmospheric phenomena that can affect remote sensing measurements are clouds as they can totally obscure a target on the surface. However, even when the sky is clear the atmosphere has a large effect on the remote sensing of water.

The gases (e.g. O_2, CO_2, O_3) and aerosols (small liquid or solid particles) present in the atmosphere absorb and scatter radiation (therefore, atmosphere also affects the other components of the total radiance). When the amount of aerosols is high the atmosphere seems hazy. This can be easily seen, for example, in a polluted city. The composition of aerosols varies in space and time and, hence, the effect they have on the radiation also varies.

The effects of this variability on water quality remote sensing can be reduced by applying atmospheric correction to the remotely sensed data. Over Case-1 waters the correction is typically done with the help of channels in the IR-region as the reflectance of water in IR is close to zero and the detected IR-radiation can be assumed to originate from the atmosphere (Gordon, 1978). Optical atmospheric models are then used to extrapolate the required correction coefficients to visible bands (i.e. to compute $T_{atm}(\lambda)$ and $L_a(\lambda)$). Finally, the corrected bands are used to retrieve water quality information (Gordon, 1978). If the radiance contribution from adjacent areas is ignored the atmospheric correction can be expressed as

$$L_u = L_d / T_{atm} - L_a. \tag{2.4}$$

Unfortunately, this method fails over Case-2 waters. The suspended particles, which are often present in Case-2 water cause a significant reflectance in the IR region. Hence, other methods are needed. These include, for example, the use of computer codes of radiative transfer equation (e.g. Second simulation of the satellite signal in the solar spectrum (6S) (Vermote et al., 1997) or Moderate Resolution Transmittance Code (Modtran) (Andersson et al., 1995)) or measuring the solar irradiance on the ground concurrently with the remote sensing measurement. The computer codes require some knowledge of the types and amounts of aerosols and gases present in the atmosphere as input parameters. These have to be measured or estimated somehow before the correction coefficients can be computed.

2.2.2 Air-water interface

The interface between the atmosphere and water (i.e. the water surface) has two effects on the radiation detected by a sensor aimed at it. First, some of the solar radiation is specularly reflected from the water surface. This can cause so called glints to appear in the image when the geometry of the measurement (the angles from the sun to the target and from the target to the sensor) is suitable. When they occur, glints usually deteriorate the quality of the data so much that the glinted areas become unusable. The radiance reflected by the water surface L_s (just above the surface) can be expressed as

$$L_s = E_{sun} R_s, \tag{2.5}$$

where E_{sun} is the solar irradiance at the surface of Earth and R_s is the reflectance of the water surface (a function of measurement geometry, wind speed and wavelength).

The second effect is that the radiation going through the interface is refracted. This means that the direction angles of the radiation change as defined by Snell's Law. This causes changes in the radiance going through the surface. These have to be accounted for in the analytical modeling of the optical signal. In other words, the radiance just above the water surface $L_{w,0}$ (without the L_s component) can be expressed as

$$L_{w,0} = L_{w,-0} \; T_s \,, \tag{2.6}$$

where $L_{w,-0}$ is the radiance just below water surface (see Chapter 2.2) and T_s is the transmittance of the water surface (also a function of the measurement geometry and wind speed).

2.2.3 Bottom effect

In areas where the water is shallow and the bottom is made from highly reflective material (e.g. sand), the reflectance from bottom can have an effect on the remotely sensed signal. In Finnish lakes the bottom often consists of dark mud and the transparency of water is generally low. Hence, the bottom effect is usually not a major problem with Finnish lakes and coastal areas.

2.2.4 Surrounding land areas (adjacency effect)

Land areas can have a high reflectance and, therefore, look bright in a remote sensing image. Even the darker land areas are usually much brighter than water. When a bright area is near to the water area measured by the instrument some of the photons reflected by the bright area are scattered into the detector by the atmosphere (shown as component 5 in Figure 2.1). This increases the apparent radiance detected from water. Some of the photons reflected by the water are scattered away from the detector but since the land area is brighter the end result is that water looks brighter than it actually is. Due to this phenomenon it may be necessary to remove the data located near land (i.e. to make a buffer area) from the analysis or to perform a correction to minimize its effect.

2.3 Theoretical basis for the optical nature of water

The spectral shape and the intensity (i.e. the color and the brightness) of the radiation emanating from water are determined by the optical properties of water and the surrounding radiation field (i.e. the light from the sun and the sky). In modeling the optical properties of water it is necessary to remove the effect the external radiation field has on the light reflected by water, i.e. to work with the inherent optical properties (IOP) of water. These are:

- Scattering coefficient $b(\lambda)$: a measure of the change of the direction of travel of photons that the radiation is composed of,
- Backscattering coefficient $b_b(\lambda)$: same as above but with a scattering angle of 90° to 180° (i.e. the direction of travel of the photon after scattering has changed more than 90°),
- Absorption coefficient $a(\lambda)$: a measure of the attenuation caused by the absorption of energy.

Studies on propagation of radiation in water based on Monte Carlo simulations and other methods (e.g. Gordon et al., 1975) have shown that the irradiance reflectance (see equation (2.2)) at just below the surface of water $R(0-,\lambda)$ can be estimated by

$$R(0-,\lambda) = \frac{E_u(\lambda,0-)}{E_d(\lambda,0-)} = f\,\frac{b_b(\lambda)}{a(\lambda)+b_b(\lambda)}, \qquad (2.7)$$

where f is a constant that includes all environmental parameters such as the zenith angle of the sun. When sun is at zenith f = 0.33 (Gordon et al., 1975). Kirk (1984) analysed the dependence of f on solar angle and found out that

$$f = 0.975 - 0.629\mu_0, \tag{2.8}$$

where μ_0 is the cosine of the solar zenith angle under water surface.

A similar equation can be written for the remote sensing reflectance above water surface (R_{RS}, equation (2.3)) (Kirk 1994):

$$R_{RS}(0,\lambda) = \frac{L_u(\lambda,0+)}{E_d(\lambda,0+)} = 0.083\ \frac{b_b(\lambda)}{a(\lambda)+b_b(\lambda)}. \tag{2.9}$$

The differences in the empirical coefficients in equations (2.7) and (2.9) are caused by the transition of radiation through the water surface and the Q-factor, which is defined as $E_u(0-)/L_u(0-)$. For most Case 2 waters, the Q-factor varies between 2.4 and 5.6 (Bukata 2005). For an ideal Lambertian surface (a surface that is perfectly diffusing) the Q factor is equal to π.

Natural water is a mixture of pure water and a variety of organic and inorganic substances originating from different sources. Each of these substances has its own wavelength-dependent scattering and absorption properties, which affect the intensity and spectral shape of the radiation. Therefore, the reflectance of water depends on the type and the amount of optically significant substances (OSS) in water and their optical properties. Inversely, the amount of OSS can be estimated with remotely sensed data if their optical properties are known. This is the basis for the remote sensing of water quality.

For natural waters four separate optically significant substances can be identified:

I. Pure water

Pure water is composed of water molecules. The scattering and absorption properties of pure water are well known and are considered constant (see e.g. Pope and Fry (1997) for absorption and Shifrin (1988) for scattering). Dissolved salt that is present in ocean water does not have absorption peaks in the visible band, and, therefore, does not affect the optical properties of water in any noteworthy way.

II. Phytoplankton

Phytoplankton is perhaps the most important optically significant substance of water as it is the base of the aquatic food chain. It is also a measure of the primary productivity of water. Phytoplankton consists of a variety of microscopic suspended organisms: phytoplankton cells (microscopic plants that use sun light for photosynthesis), and bacteria and viruses that covary with it.

In a phytoplankton cell the absorption of solar radiation takes place in photosynthetic pigments, of which the most commonly analyzed by the remote sensing community is chlorophyll *a* (chl *a*) as its concentration can be used as a proxy for phytoplankton biomass. Chl *a*, and consequently the phytoplankton cell, has two absorption peaks at around 430 nm and 665 nm (blue and red, hence the green color) (Mobley, 1994). However, the magnitude and the exact position and the width of these peaks depend on the cell species and its physiological state (Kirk, 1983).

The scattering coefficient of phytoplankton is quite low (Bricaud et al., 1983). Therefore, the covarying viruses and other small organisms are believed to contribute

to the backscattering that is observed in optical measurements of natural water (Ulloa et al., 1992).

Phytoplankton also has fluorescence properties with a peak centered at about 685 nm. This property can also be used to monitor phytoplankton with fluorometers (Yentsch and Yentsch, 1979).

In a laboratory, the concentration of chl *a* can be measured by first extracting the chl *a* pigments from the phytoplankton cells by using e.g. hot ethanol. The sample is then filtered and the filter is examined with a spectrophotometer. Next the filter is bleached with acid and the filter is examined again with a spectrophotometer. The concentration of chl *a* is then determined from the two spectral measurements.

III. Suspended inorganic material

Suspended inorganic material is, as the name states, the suspended matter that does not fall into the organic phytoplankton component. It includes sand, clay and other bottom sediments, runoff material from rivers and also dust from atmospheric sources. Suspended inorganic material is usually found only relatively near the shore, so in most cases only coastal and inland waters are affected by it.

The optical properties of suspended inorganic material depend on the geology of the target area. For example, the shape and the size distribution of particles, have a major impact on the scattering and absorption coefficients (Bukata et al., 1995). As the type of the geological material (and therefore, the shape and the size distribution of particles) varies from one location to another, the optical properties of suspended inorganic material determined for one area are often not valid for another target. This complicates optical modeling and hampers the development of water quality algorithms.

The concentration of suspended inorganic material can be determined gravimetrically. First, a sample of water is filtered through a filter, which retains all particles above a certain size (e.g. 0.45 μm). The filter is then dried to remove any water and combusted to remove any organic matter. The difference in the weight of the filter before filtering and after combustion is the amount of inorganic matter. However, due to the complexity of the method it is more common (at least in Finland) to measure total suspended solids (TSS). It is also measured gravimetrically before the combustion. Hence, TSS includes both inorganic and organic matter.

IV. Colored dissolved organic matter (CDOM)

CDOM is a class of substances consisting of various humic and fulvic acids originating from, for example, the degradation of phytoplankton (Kirk, 1983). Other sources for CDOM are rivers running through forests and swamps. CDOM has also been called yellow substances, humus, gelbstoff or gilvin.

The absorption spectrum of CDOM has been found to follow an exponential function in the visible band (high absorption values with shorter wavelengths; very little absorption beyond 700 nm) (Bricaud et al., 1981). As the substances are assumed to be fully dissolved (i.e. it does not contain particles that cause scattering), the scattering properties of CDOM can be ignored in remote sensing. CDOM usually includes the detrital component that has optical properties so close to those of CDOM that it is difficult to separate the two.

CDOM can be determined from a water sample by using e.g. spectrophotometry. It is usually expressed as an absorption coefficient at a certain wavelength (e.g. 400 nm, hence $a_{CDOM}(400)$ with unit m^{-1}), not as concentration like chl *a* and TSS.

Total absorption and scattering coefficients

The individual absorption and scattering properties of the four substances discussed above can be used to derive the total absorption and backscattering coefficients $a_{tot}(\lambda)$ and $b_{b,tot}(\lambda)$. The coefficients can be expressed as a sum of the individual components:

$$a_{tot}(\lambda) = a_w(\lambda) + a_y(\lambda) + C_{ph}\, a_{ph}(\lambda) + C_s\, a_s(\lambda) \tag{2.10}$$

and

$$b_{b,tot}(\lambda) = b_w(\lambda) + C_{ph}\, b_{ph}(\lambda) + C_s\, b_s(\lambda), \tag{2.11}$$

where C_{ph} and C_s are the concentrations of phytoplankton and suspended inorganic matter, respectively, and the subscripts *w, y, ph* and *s* in *a* and *b* stand for water, CDOM, phytoplankton and suspended inorganic matter. The backscattering coefficient of phytoplankton is low (Bricaud et al. 1983) and hence usually neglected.

Equations (2.7), (2.10) and (2.11) can be used to form a bio-optical reflectance (BOR) model. BOR-models use OSS values as input and compute the corresponding reflectance above or under water surface. A bio-optical model parameterized for Finnish lakes has been developed in Kallio et al. (2005) and Kallio (2006) (see Appendix A for details). Sample spectra derived with this model are presented in Figure 2.3. The spectra are simulated by allowing one OSS to change while the others are kept constant (with value = 1). This procedure shows the theoretical effect of each OSS (separately from the rest) on the spectra when compared with the baseline case (values of all OSS are equal to 1). In nature the values of different OSS follow each other (see Table 1.5 for correlations) and the spectral features are combined. For example, according to the model the reflectance decreases when the chl *a* concentration increases (chl *a* is assumed to only absorb in the model). In a real lake, the increasing chl *a* would always be accompanied with increasing TSS and hence the reflectance would also increase.

Figure 2.4 shows sample spectra from two open ocean areas (Case 1 water) and one turbid coastal area (Case 2 water). The spectral differences between the areas are clear. The scale of Figure 2.4 is logarithmic; hence, a direct comparison with the spectra in Figure 2.3 is difficult. Nevertheless, the shape of the coastal case is similar to the spectra presented in Figure 2.3a (chl *a* is allowed to change). As mentioned earlier, Finnish lakes and coastal areas are all assumed to belong to Case 2 waters.

The spectral features shown in Figure 2.3 and Figure 2.4 can be used in the selection of the channels when the retrieval algorithms are defined. For example, in Case 1 waters the channels for chl *a* estimation are usually selected from the blue and green areas of the spectrum (440 nm - 550 nm) where the signal from chl a is strong (at about 440 nm chl *a* absorbs strongly while at 550 nm the absorption is low). In Case 2 water, the absorption by CDOM makes these channels unreliable as shown in Figure 2.3c (absorption increases towards the shorter wavelengths).

Fortunately, chl *a* also affects other areas of the spectrum and these can be used for its estimation in Case 2 areas instead of the blue-green wavelengths. For example, Figure 2.3a shows that the peak at about 700 nm grows as the concentration of chl *a* increases. It is due to the local absorption maximum of chl *a* (at about 665 nm),

absorption by pure water and backscattering by particles (e.g. Gitelson 1992, Gitelson et al. 2000). Near 700 nm the effects of CDOM are not as strong as in the blue-green wavelengths.

Another feature visible in Figure 2.3 is that the increasing TSS-concentration causes increasing reflectance in all wavelengths between 400 and 700 nm.

2.4 Other water quality variables

Although the substances present in water are usually divided into the three groups presented above some other water quality variables that can be measured with remote sensing instrument are also worth discussing. These are often a measure of the combined effect of the optically significant substances and their measurement is sometimes more subjective.

Secchi depth

Secchi depth is a measure of the transparency of water (Lindell et al., 1999). Thus, all OSS presented above affect it (see Table 1.5 for correlations). If the amount of chl *a*, TSS and CDOM is low, Secchi depth is large. If the amounts of OSS are high, Secchi depth is low.

Its popularity is probably based on the ease of the measurement: a standard size white (or black and white sectors) disk (diameter 12 to 20 cm) is lowered into water and the depth where the disk can no longer be separated from the surrounding water is defined as the Secchi depth. The measurement depends on the lighting conditions and on the vision of the observer and, therefore, is somewhat subjective.

From the biological point of view Secchi depth is important because it is a measure of the water depth into which sunlight penetrates thus defining the thickness of the water layer, where sunlight is available for photosynthesis.

Turbidity

Turbidity is another commonly used water quality variable. It is a measure of optical scattering in the water and, hence, closely related to the amount of suspended particles. I.e. if the amount of chl *a* (phytoplankton) and/or suspended sediments is high, the turbidity value is also high.

Remote sensing instruments can usually measure turbidity with good accuracy. Typically, a single band in the visible or near-IR region can be used to map turbidity with reasonable accuracy (Lindell et al., 1999).

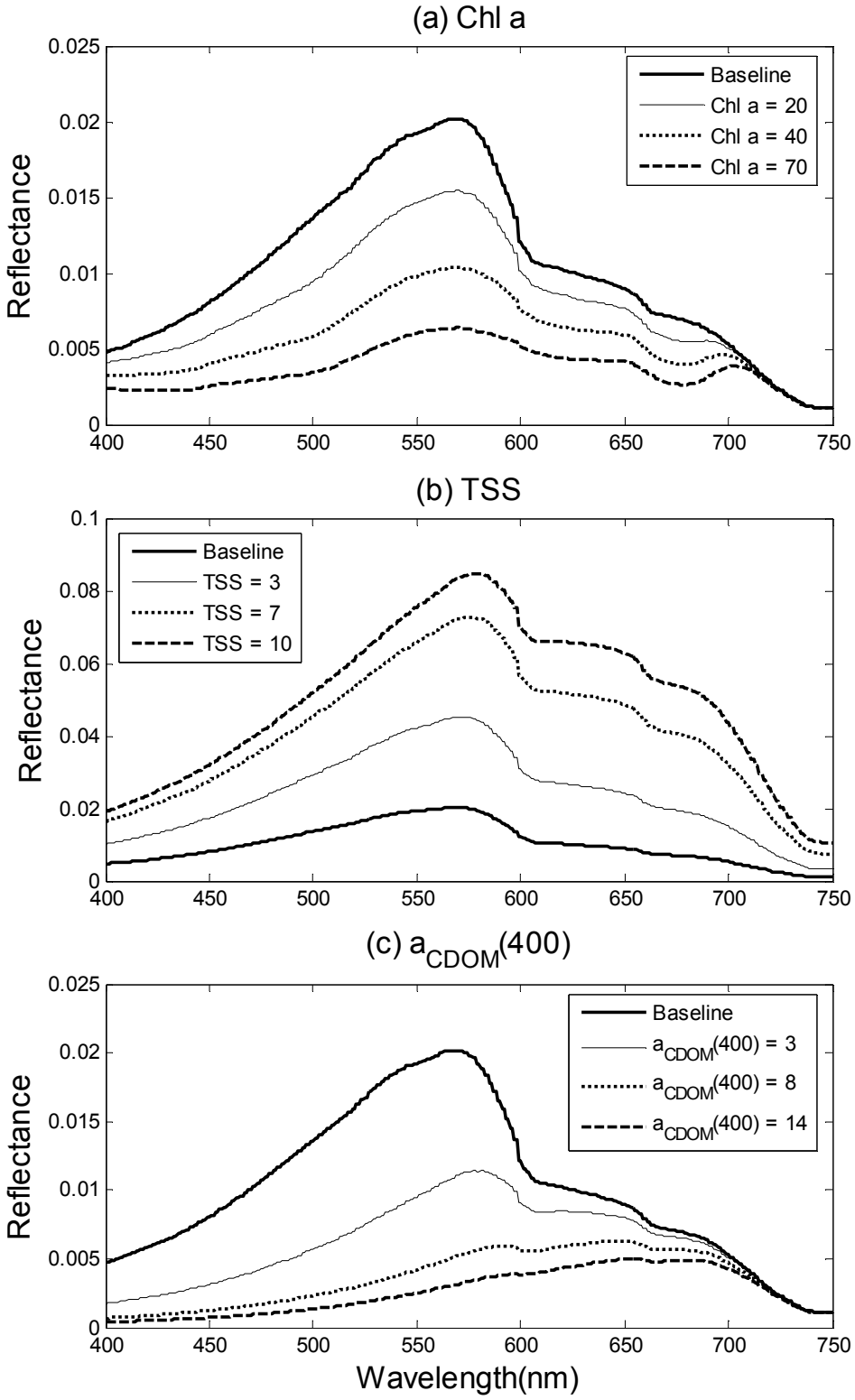


Figure 2.3. Underwater spectra computed with a bio-optical model developed for Finnish lakes (see Appendix A; Kallio et al. (2005) and Kallio (2006)). (a) Chl a has values 20, 40 and 70 μg/l while TSS = 1 mg/l and $a_{CDOM}(400) = 1\ m^{-1}$. (b) TSS has values 3, 7 and 10 mg/l while chl a = 1 μg/l and $a_{CDOM}(400) = 1\ m^{-1}$. (c) $a_{CDOM}(400)$

has values 3, 8 and 14 m^{-1} while chl a = 1 μg/l and TSS = 1 mg/l. The baseline in each plot is computed using values chl a = 1 μg/l, TSS = 1 mg/l and $a_{CDOM}(400) = 1$ m^{-1}.

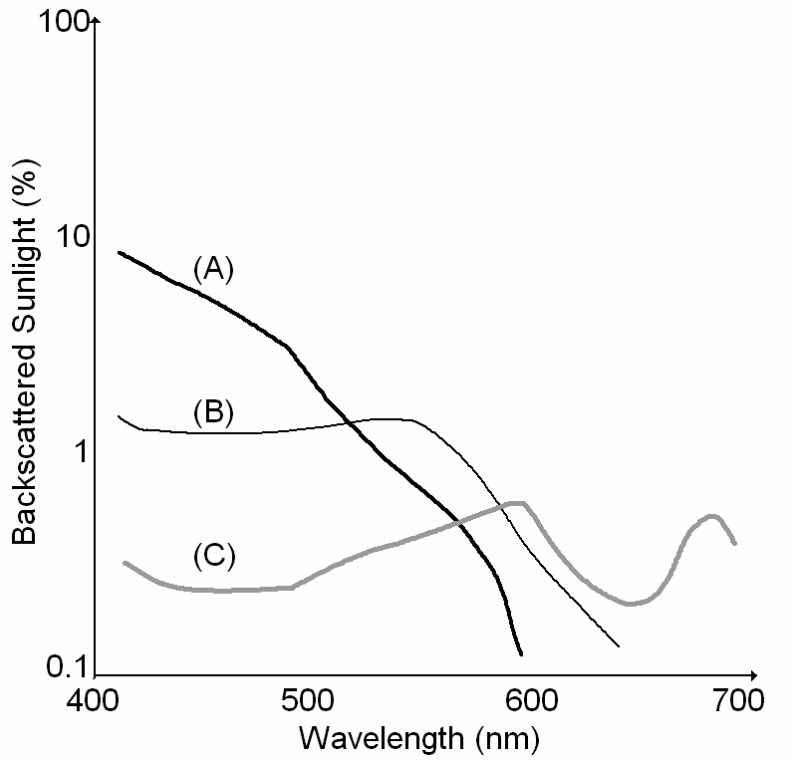


Figure 2.4. Percentage of sunlight backscattered from upper ocean layers as a function of wavelength in nanometers, under three conditions: (A) clear open ocean water, low phytoplankton concentration (Case 1); (B) moderate phytoplankton bloom, open ocean (Case 1); (C) turbid coastal waters containing sediment as well as phytoplankton (Case 2). [Redrawn from NASA (2006), with permission]

3 Advances in remote sensing of Case 2 waters

Due to the reasons presented in previous chapters it is often difficult to compare remote sensing results obtained from different geographical areas. Nevertheless, an attempt is made here to review the previous done work on remote sensing of lake and coastal water quality. The objective is to find out the current state of the art of remote sensing of water; hence, the review will concentrate on the more recent (during years 2000-2005) results obtained with instruments that are still in use instead of starting from the early days. The review includes only those publications where real (i.e. not simulated) remote sensing data (from aircraft and satellites, not boats or other ground based sources) have been used for water quality monitoring so that it is possible to compare them with the results of this study. Also, studies where *in situ* data analyzed in a laboratory have not been used are not included in this analysis.

3.1 Lakes

A summary on the state-of-the-art methods for lakes is presented in Table 3.1. The largest collection of recent remote sensing results relevant to lake monitoring in Europe is available in the thematic issue of the journal "the Science of the Total Environment" published in March 2001. In the thematic issue, Giardino et al. (2001) obtained a good correlation between atmospherically corrected Landsat TM reflectances and *in situ* chl *a* (the coefficient of determination $R^2 = 0.99$) and Secchi depth ($R^2 = 0.85$) for Lake Iseo, Italy. However, they only had four data points and the variability of the *in situ* data was limited (from 5 to 8 μg/l for chl *a* and 4–7 m for Secchi depth). For chl *a* they used a channel difference (TM1 - TM2) algorithm and for Secchi depth channel ratio (TM1/TM2).

Kallio et al. (2001) used AISA airborne spectrometer data to test and develop algorithms (including single channels, channel ratios and channel differences) for five water quality variables (chl *a*, turbidity, total suspended solids, absorption coefficient of aquatic humus, and Secchi depth) for 11 lakes in Finland. The number of data points varied from 47 to 105, depending on water quality variable and R^2 ranged from 0.84 to 0.95 after the data were grouped seasonally.

Pulliainen et al. (2001) also used AISA data for investigating monitoring of chl *a* in Finland. Their results showed that, when remote sensing data are used with *in situ* data, the error of the chl *a* estimate is smaller than when only *in situ* data are used. They also showed that humic lakes, for which the chl *a* algorithm fail, can be discriminated by analyzing the shape of the spectrum measured for a lake.

Östlund et al. (2001) used Landsat-TM and CASI airborne spectrometer data for measuring chl *a* (including phaeophynite *a*) and suspended particulate matter (SPM) for Lake Erken, Sweden. The results with TM data were good, with R^2 over or close to 0.9 for several algorithms (e.g. chromaticity green or blue and log transformed channel ratios). However, they only had six data points. Furthermore, the variability of TM-data was low (only a few digital numbers over the whole Lake Erken). That and the striping present in TM-data restrict the usability of TM-data for this application. When airborne data were compared with the chl *a* data derived from a flow-through fluorometer the highest R^2 was 0.84 (single channel at 659.0 - 669.8 nm). For SPM the best R^2 was 0.82 (same channel as for chl *a*). With a more conventional method of relating CASI data to samples analyzed in laboratory (with

nine training data points and ten testing data points) the lowest RMS error obtained was 6.25 μg/l for chl *a* and 0.57 mg/l for SPM (single channel at 700.2 - 709.1 nm). The variability ranges of the test data were 2.9 - 48.1 μg/l for chl *a* and 1.7 - 7.1 mg/l for SMP.

Dekker et al. (2001) took the analytical modeling approach and used Landsat-TM and SPOT images (acquired between May and August 1995), a large spectral library (data collected between May and September 1995) and a bio-optical model, which was used to predict water quality over one year, to map the total suspended matter in Frisian lakes in the Netherlands. The *in situ* TSM data were derived Secchi depth and chl *a* data. The accuracy of the method was not, however, reported as a numeric value.

George and Malthus (2001) used an airborne spectrographic imager (CASI) to map chl *a* in five lakes in Scotland. They reported R^2 of 0.94 for their best algorithm (15 data points). They also concluded that the blue-green ratio used in the ocean applications also works for inland water if the concentration of optically significant substances is low (in their case the mean chl *a* concentration was 5 μg/l and the maximum was 10 μg/l).

In other publications Thiemann and Kaufmann (2000a) used LISS-III (Linear Imaging Self Scanning System) data from the Indian satellite IRS-1C (resolution of multispectral data is 23.5 m) and linear spectral unmixing to retrieve chl *a* values and trophic state of about 60 lakes in Germany. Linear spectral unmixing is a method where the reflectance of each pixel (and in each band) is assumed to be a linear combination of the spectral characteristics of the material present in the pixel. The spectrum information of the pixel can thus be used to retrieve the proportion of the material. For developing the algorithm they used 11 *in situ* data points from five lakes (however, one of the points is fairly dominating). They obtained R^2 of 0.85 and a mean standard error of 13 μg/l for chl *a*. The variability range of the data was 2 - 70 μg/l.

Thiemann and Kaufmann (2000b, 2002) used data from two airborne hyperspectral spectrometers, CASI and HyMap, to estimate chl *a* and Secchi depth levels of 31 lakes in Germany during 1997-1999. For Secchi depth determination they used the normalized mean of the spectrum between 400 nm and 750 nm as an index and an exponential function between the index and *in situ* data. That yielded R^2 of 0.85 and mean standard error of 1.2 - 1.3 m. For chl *a* they used channel ratio (678 nm / 705 nm) and got R^2 of 0.89 and mean standard error of 10.2 – 10.9 μg/l. The number of data points (from nine lakes and five dates) they used was about 30 (Thiemann, 2001).

Strömbeck et al. (2003) used a band-ratio algorithm (620 nm/705 nm) with atmospherically corrected MERIS data to estimate the chl *a* values in Lake Garda (Italy). The *in situ* data were collected with a flow-through system (over 3000 samples). These samples were averaged so that they match the MERIS pixels. The chl *a* concentrations in the data were very low and the variability is small (0.7-2.6 μg/l).

Vincent et al. (2004) used Landsat-TM data for mapping turbidity in Lake Erie (USA). They used a ratio of TM-channels 3 and 2, which gave R^2 of 85.2 % with 30 samples (from a single day).

From the results listed above we can conclude that airborne sensors have been popular in lake monitoring due to their capability to measure targets with high spatial and spectral resolution. Other commonly used instruments are high-resolution satellite

spectrometers (e.g. Landsat-TM and SPOT). While the parameters of individual lakes or even small lake groups can be measured with good accuracy (R^2 around 0.9) the algorithms or methods (typically channel ratios) that yield the best correlation vary. Also, the data collected for the studies are often from a short time period and include a limited amount of data points with limited variability in the water quality parameters or in the remotely sensed data. Therefore, their applicability in different locations and during different seasons may be limited. The usability of these results for comparison with the results presented here is also limited and it may only be possible to perform qualitative analysis.

3.2 Coastal areas

A summary on the state-of-the-art methods for coastal areas is presented in Table 3.2. With coastal and oceanic areas the spaceborne instruments, such as SeaWiFS, MODIS and MERIS, are more popular. This is due to the need to cover larger areas. That is possible only with satellite data since airborne monitoring of large areas is expensive. Also, in most coastal areas the spatial resolution of the instrument is not as critical factor as with lakes.

The operators of satellite instruments (such as ESA and NASA) often provide their own products (Level 2 or higher) that include values for parameters such as chl *a*. These products are typically derived with global algorithms using complex methods such as neural networks. The problem with these global algorithms is that they can fail when the scale is something smaller than global. This seems to be especially true for the Baltic Sea. For example, in a study by Gregg and Casey (2004), where SeaWiFS was used for estimating chl *a*, the global R^2 value was 0.6 (2479 data points), however, when the data for Baltic and North Seas were analyzed the R^2 was 0.02 (34 data points). The likely reason for the failure is that the water in the Baltic Sea is so different from those waters the global algorithms were developed for. Baltic Sea is quite turbid and contains a lot of CDOM. These factors also complicate the atmospheric correction of satellite data. By comparing the range of values presented in Table 3.2 to those presented in Table 1.4 the differences between Baltic Sea and other areas becomes clear. For example, the maximum values of chl *a* and CDOM are clearly higher in the Baltic.

Darecki and Stramski (2004) compared the results of the standard MODIS chl *a* and $a_{CDOM}(400)$ algorithms against *in situ* observations in the Baltic Sea and observed large errors. The smallest RMS error for chl *a* was 157 % (with an algorithm called *chlor_a_3*); however, the data range is quite small. For $a_{CDOM}(400)$ the RMS error is smaller (21 %). The large errors in chl *a* estimation were believed to originate from inadequate in-water bio-optical algorithms and unsuccessful atmospheric correction.

In other studies, Jorgensen (2004) used a double ratio algorithm with SeaWiFS data in the Danish waters (North Sea and Baltic Sea). This algorithm gave better results (R^2 = 0.54) than the standard (OC4v4) SeaWiFS method (R^2 = 0.39). The variability range of the data was quite small (0.2-7.5 µg/l).

D'Alimonte and Zibordi (2003) used a neural network method for estimating chl a in the Adriatic Sea (Italy). They first developed the algorithm with *in situ* data that include chl *a* concentrations (variability range 0.1-9 µg/l) and in-water measurements of upwelling radiance. The algorithm was then tested with atmospherically corrected SeaWiFS data. The resulting R^2 value was 0.8 (31 data points). The data include both

Case 1 and Case 2 areas and the authors did not separate them for the SeaWiFS analysis. Hence, the applicability of the method for Case 2 water only is not known.

D'Sa and Miller (2003) used SeaWiFS data to estimate chl *a* in the Gulf of Mexico, near the Mississippi River. They used an empirical band-ratio algorithm (490 nm/555 nm) calibrated with in-water optical measurements and tested it with SeaWiFS observations (6 data points, variability range 2-12 µg/l). They also used the OC4 algorithm in order to test the accuracy of the standard method. The results show that the empirical algorithm performs slightly better, although it still underestimates the chl *a* values.

Gohin et al. (2002) took the standard OC4 SeaWiFS algorithm and added two more channels (412 nm and 555 nm) into it. The proposed algorithm (named OC5) reduces the over-estimation error that the OC2 and OC4 algorithms often have in coastal areas. New parameters for the algorithm are derived using data from the Bay of Biscay (France) and the English Channel (United Kingdom) and the resulting R^2 value is 0.7 (178 data points, variability range 0.2-44 µg/l).

Hu et al. (2004) used medium resolution (250 m and 500 m) MODIS data to map chl *a*, $a_{CDOM}(400)$ and TSS values in Tampa Bay (Florida, USA). The *in situ* data were collected over two days using a flow-through device. For TSS they used a channel difference algorithm: R(645)-R(859). The algorithm is based on the assumption that the water leaving radiance at 859 nm is zero and the detected radiance is caused by the atmosphere. The resulting $R^2 = 0.9$ (31 data points, variability range 1-11 mg/l).

For TSS the authors reported results for only one measurement day. For chl *a* and $a_{CDOM}(400)$ data they presented result for two single day cases and for a case when both data sets were combined. The algorithms for $a_{CDOM}(400)$ and chl *a* were different for both separate days and for the combined data set and they were based on complex ratios of channel differences. The R^2 values for chl *a* were 0.93 (n = 95), 0.16 (n = 31), and 0.72 (n = 126) for day 1, day 2, and the combined set, respectively. For $a_{CDOM}(400)$ the R^2 values were 0.59 (n = 95), 0.80 (n = 31), and 0.72 (n = 126).

The 250 m MODIS data were also used in Miller and McKee (2004) to estimate water quality Northern Gulf of Mexico (USA). They used a simple atmospheric correction (dark pixel subtraction) to process data obtained from six field campaign. The corrected reflectances were then used to estimate TSM. The R^2 value was 0.89 (52 data points, variability range 0-56 mg/l).

Zibordi et al. (2003) compared the MERIS level 2 product with *in situ* chl *a* and TSM data in the Adriatic Sea (Italy). They observed an over-estimate by a factor of 2 in the satellite chl *a* data (Algal I). The variability range of *in situ* data was 0.2-4.1 ug/l. When a regional band ratio algorithm (490 nm/555 nm) is applied to the data the over estimation is reduced and the R^2 value increases to 0.72 (from 0.69). For TSM data the level 2 MERIS product did not show any correlation. The application of regional algorithm improved the results; however, the authors did not present any numeric values of the accuracy. The variability range of *in situ* TSM data was 0.1- 3 mg/l.

Table 3.1. State-of-the-art of the monitoring of water quality parameters in lakes. The Method column shows the algorithm (wavelengths if possible) used in the study, N is the number of data points, R^2 is the coefficient of determination, Error is usually the RMSE[(1)] or standard error[(2)] in absolute units or as a relative (percentage) value (usually divided by the mean estimated value of the parameter in question), Data range is the maximum and the minimum value of the investigated water quality parameter, Lake N is the number of lakes.

Parameter	Instrument	Method	N	R^2	Error	Data range	Lake N	Area	Source
Chl *a*	MERIS	R(620)/R(705)	12	0.76	0.27 µg/l	0.7-2.6 µg/l	1	Italy	Strömbeck et al. (2003)
Chl *a*	CASI	R(705)/R(678)	30	-	10.2 µg/l	1-100 µg/l	31	Germany	Thiemann and Kaufmann (2002)
Chl *a*	Landsat-TM	TM1-TM2	4	0.99	0.05 µg/l	5.5-8 µg/l	1	Italy	Giardino et al. (2001)
Chl *a*	CASI	R(705)/R(681)	-	0.47	-	4-39 µg/l	1	Sweden	Strömbeck et al. (2001)
Chl *a*	Landsat / CASI	R(704)	19	0.97	8.52 µg/l	2-78 µg/l	1	Sweden	Östlund et al. (2001)
Chl *a*	AISA	L(702)/L(673)	88	0.91	6.03 µg/l	2-100 µg/l	9	Finland	Kallio et al. (2001)
Chl *a*	CASI	R(704)	19	0.97	6.25 µg/l	2.9-50.6 µg/l	1	Sweden	Östlund et al. (2001)
Chl *a*	Landsat-TM	Chromaticity green	6	0.93		2.1-27.4 µg/l	1	Sweden	Östlund et al. (2001)
Chl *a*	CASI	Blue/green	15	0.94		0.5-10 µg/l	5	Scotland	George and Malthus (2001)
Chl *a*	CASI/Hymap	R(705)/R(678)	30	0.89	10.9 µg/l	1-100 µg/l	9 (31)	Germany	Thiemann and Kaufmann (2000b)
Chl *a*	IRS-1C LISS-III	Linear spectral unmixing	11	0.85	13 µg/l	2-70 µg/l	5 (60)	Germany	Thiemann and Kaufmann (2000a)
Secchi	CASI	See the source	-	-	1.2 m	0.25-8.5 m	31	Germany	Thiemann and Kaufmann (2002)
Secchi	Landsat-TM	TM1/TM2	4	0.85	0.45 m	4.5-7 m	1	Italy	Giardino et al. (2001)
Secchi	AISA	(L(492)-L(751))/(L(621)-L(751))	103	0.86	0.51 m	0.4-7.0 m	10	Finland	Kallio et al. (2001)
Secchi	CASI/Hymap	See the source	-	0.85	1.2 m	~1-5 m	5 (31)	Germany	Thiemann and Kaufmann (2000b)
Secchi	Landsat-TM	TM1/TM2	4	0.85		4.6-6.8 m	1	Italy	Giardino et al. (2001)
SM[(3)]	Landsat-TM	(TM2 + TM3)/2	-	-	-	7-92 mg/l	6	Holland	Dekker et al. (2001)
SM	CASI	R(704)	19	0.82	0.57 mg/l	1.7-7.3 mg/l	1	Sweden	Östlund et al. (2001)
SM	Landsat-TM	TM1	6	0.95		1.45-5.25 mg/l	1	Sweden	Östlund et al. (2001)
Turbidity	ETM+ / TM	TM3/TM2	30	0.85	1.58 NTU	1-14 NTU	1	USA	Vincent et al. (2004)
Turbidity	AISA	R(709)	105	0.93	1.9 FNU	0.4-26 FNU	11	Finland	Kallio et al. (2001)

(1) See Eq. (4.3)

(2) Standard error = $\sigma/n^{1/2}$, where σ is the standard deviation and n is the number of samples.

(3) SM = Suspended matter

Table 3.2. State-of-the-art of the monitoring of water quality parameters in coastal areas. The Method column shows the algorithm (wavelengths if possible) of used in the study, N is the number of data points, R^2 is the coefficient of determination, Error is usually the standard error in absolute or a relative (percentage) value of the mean value of the parameter in question (see Table 3.1 for more details), Data range is the maximum and the minimum value of the investigated water quality parameter.

Parameter	Instrument	Method	N	R^2	Error	Data range	Area	Source
Chl *a*	SeaWiFS	Level 3 product	2479	0.60	-	0.02-50 µg/l	Global [2]	Gregg and Casey (2004)
Chl *a*	SeaWiFS	Level 3 product [1]	34	0.02	49.4 % RMS log Error	-	Baltic/North Seas [2]	Gregg and Casey (2004)
Chl *a*	MODIS	See the source	126	0.72	-	11-22 µg/l	USA [3]	Hu et al. (2004)
Chl *a*	MODIS	L(488)/L(551)}	12	-	157 %	0.6-3 µg/l	Baltic Sea	Darecki and Stramski (2004)
Chl *a*	SeaWiFS	[R(510)/R(555)]/[R(443)/R(670)]	28	0.54		0.2-7.5 µg/l	Denmark	Jorgensen (2004)
Chl *a*	SeaWiFS	Neural network	31	0.80	30 %	0.1-9 µg/l	Italy	D'Alimonte and Zibordi (2003)
Chl *a*	SeaWiFS	R(443)/R(550)	55	0.937	-	0.05-0.45 µg/l	Israel	Iluz et al. (2003)
Chl *a*	MERIS	R(490)/R(555)	28	0.72	-	0.2-4.1 µg/l	Italy	Zibordi et al. (2003)
Chl *a*	SeaWiFS	R(490)/R(555)	35	0.71	40 %	0.2-5 µg/l	Italy	Mélin et al. (2003)
Chl *a*	SeaWiFS	R(490)/R(555)	6	0.88	-	2-12 µg/l	USA [3]	D'Sa and Miller (2003)
Chl *a*	SeaWiFS	OC4	13	-	-	0.1-0.7 µg/l	China	Tang et al. (2003)
Chl *a*	SeaWiFS	OC2, OC4, OC5	178	0.70	0.66 (66 % of the median)	0.2-44 µg/l	France/UK	Gohin et al. (2002)
Chl *a*	SeaWiFS	See the source	16	-	18 %	0.5-3 µg/l	China [2]	He et al. (2000)
SM[4]	MODIS	R(645) – R(859)	31	0.9	-	1-11 mg/l	USA [3]	Hu et al. (2004)
SM	MODIS	R(645)	52	0.89	4.74 mg/l	0-55 mg/l	USA [3]	Miller and McKee (2004)
SM	MERIS	R(708)	5	-	35 %	-	Belgium	Nechad et al. (2003)
SM	SeaWiFS	R(510)/R(670)	32	0.75	28 %	0.3-4 mg/l	Italy	Mélin et al. (2003)
CDOM	MODIS	See the source	9	-	21 %	0.4-1 m^{-1}	Baltic Sea	Darecki and Stramski (2004)
CDOM	MODIS	See the source	31	0.80	-	2-2.5 m^{-1}	USA [3]	Hu et al. (2004)
CDOM	MODIS	See the source	126	0.72	-	0.9-2.5 m^{-1}	USA [3]	Hu et al. (2004)

[1] The study is the same as immediately above, except for the target area.
[2] The results include areas that are not coastal by definition.
[3] Also includes river delta data.
[4] SM = Suspended matter.

4 Development of retrieval methods for Finnish Case 2 waters

This chapter presents the research performed in [P1]-[P6]. First the objectives of the work are discussed and the *in situ* and remote sensing data used in the analysis are presented. Then the methods used in the analysis are explained and the results obtained with them are presented. The methods and results are presented for two themes. The first is the estimation of water quality parameters using empirical methods [P1]-[P5]. The second is classification of lakes using remote sensing data [P5]-[P6].

4.1 Objectives

The greatest problem in remote sensing of lakes (and coastal areas) is that a globally applicable method for estimating water quality parameters has not yet been developed for them. The absorption and scattering characteristics of OSS in one lake can be different from those in another and a method developed for one lake can give erroneous results in the other. Especially, the amount of CDOM is in Finnish lakes generally higher than in other European lakes. Thus, the lakes in Finland require retrieval methods developed specifically for them.

With that in mind, the overall objectives of this thesis are to examine the use of airborne and spaceborne remote sensing data for monitoring water quality in Finnish lakes and coastal areas and to develop methods, which improve the operational status water monitoring. This will allow a greater portion of the lakes (both the number of lakes and the area covered by them) and coastal areas to be monitored more often and will lead to improved knowledge of the status natural waters in Finland. Hence, the Finnish water monitoring programs will be in better compliance with EU directives.

The specific objectives are:

1. To develop and test algorithms for the estimation of water quality parameters in Finnish conditions using airborne and spaceborne remote sensing data
2. To test lake classification with remote sensing data
3. To compare the results obtained here with results from other publications.

4.2 Data

The data used in this work were obtained from the following sources:

1. The Finnish Environment Institute (SYKE) and local environmental centers routinely collect and analyze water samples from about 5700 lake and 1310 coastal sampling stations. All data from local, regional and national monitoring programs are stored in a database called the Environment Register. Samples from this data were used to quantify the relationship between MODIS data and *in situ* water quality. These data were used in [P4].
2. From time to time, the data collected to the Environment Register are used for classifying lakes, rivers and coastal areas into five quality classes ranging from excellent to poor. The final result of the classification is a map and the version that was published in 1999 is used as ground truth for MODIS data. These data were used in [P6]

3. An extensive set of concurrent (the time difference was usually less than 3 hours) airborne and *in situ* data were collected in four water quality measurement campaigns during 1996-1998. The campaigns were part of the EU-funded SALMON project (Lindell et al., 1999) and national remote sensing projects. They were conducted by the Laboratory of Space Technology of the Helsinki University of Technology (LST) and SYKE. These data were used in [P1], [P5], and [P2]
4. On April 27, 2004 LST and SYKE, in co-operation with the Luode Consulting Oy, performed a measurement campaign near the City of Helsinki (near Vuosaari harbor, coastal data) using boat- and airborne instruments. These data were used in [P3]
5. MODIS data were obtained from NASA. Due to the generous data policy of NASA the scientific community can use MODIS data free of charge. The MODIS data used in this work were acquired on August 27, 2000 at 10:00 GMT. These data were used in [P4] and [P6].
6. MERIS data were obtained from ESA under the AO program. The MERIS data used in this work were acquired April 27, 2004 at 09:35 GMT. These data were used in [P3].

4.2.1 *In situ* data

SALMON project

The *in situ* measurements performed in Finland during the SALMON project included water sampling for laboratory analysis (e.g. chlorophyll *a*, turbidity, total suspended solids, aquatic humus), on-site measurements (e.g. Secchi depth, upwelling and downwelling irradiance with an underwater spectrometer) and weather observations (e.g. wind speed and direction, cloudiness). The sum of chlorophyll *a* and phaeophytin *a* (denoted here with chl *a*) was determined with a spectrophotometer after extraction with hot ethanol (ISO 10260) and turbidity by the nephelometric method (based on the measurement of light (860 nm) scattered within a 90° angle from the beam directed at the water sample, ISO 7027). In total phosphorous determination the water sample was digested by potassium peroxodisulphate before analysis with ammonium molybdate (Murphy and Riley, 1962).

The number of measurement days was eight (see [P1] Table 2 for a list of dates) and the number of lakes was 11. The data set covers all major lake types encountered in Finland (from oligotrophic to eutrophic, and humic; see Chapter 1.1) and the number of usable (i.e. non-cloudy) data points with concurrent spectrometer and *in situ* data was 122. Other results from this project have been presented by Härmä et al. (2001), Kallio et al. (2001), and Pulliainen et al. (2001).

SYKE routine monitoring program

The water quality variables analyzed in the routine monitoring programs include color (aquatic humus), Secchi depth, turbidity, chlorophyll *a*, hygienic indicator bacteria, phosphorous, O_2 saturation and depletion and several poisonous substances. The data are used for classifying lakes into five classes roughly every four years. For the classification used in [P6] the data were collected during 1994-1997, included 2.5 million analysis results and the analysis itself took two more years (SYKE, 2006). The classification is based on a statistical analysis of the collected water quality

parameter data. The final decision will be made by a water quality expert. Thus, the classification result is somewhat subjective.

The frequency of sampling varies from station to station. Some are sampled only a couple of times during the ice-free period while the intense monitoring stations are visited roughly every two weeks. The number of water quality variables analyzed from each station also varies. At the time of writing [P4], the available number of sampling stations for which water samples were collected within plus/minus three days of the MODIS overpass on August 27, 2000 was 134. Of those turbidity was analyzed for 116 stations. The analysis methods are the same as those used during the SALMON campaigns.

Vuosaari-campaign

The *in situ* data collected during the Vuosaari-campaign (coastal site near the city of Helsinki on April 27, 2004) include a 28 km long transect measured with a flow-through measurement system (Lindfors et al. 2005) installed on a boat (measurement depth 0.5 m; 5103 data points), water samples collected at ten fixed stations along the transect (measurement depth 0.5 m), and five Secchi depth (SD) measurements. The samples were analyzed in a laboratory and yielded concentrations for chl *a* (the sum of chlorophyll *a* and phaeophytin *a*) and TSS, and the absorption coefficient of CDOM at 400 nm ($a_{CDOM}(400)$).

Chl *a* was determined with a spectrophotometer after extraction with hot ethanol (ISO 10260). TSS was determined with the gravimetric method using a GF/F filter. $a_{CDOM}(400)$ was measured with a spectrophotometer from a sample filtered through a GF/F filter.

The flow-through system measures temperature, conductivity, and total absorption and scattering coefficients ($b_{tot}(\lambda)$ and $a_{tot}(\lambda)$, respectively) at 9 wavelengths between 412 and 715 nm using the WET Labs ac-9 absorption and attenuation meter. $b_{tot}(\lambda)$ and $a_{tot}(\lambda)$ are transformed into OSS values (chl *a*, TSS and $a_{CDOM}(400)$) using empirical algorithms developed with water sample data [P3]. This yields 5103 data points with full OSS information. The resulting dataset can be used with remote sensing data for further development and testing of retrieval algorithms.

The flow-through data were also pre-processed by finding the flow-through OSS values located within each MERIS pixel and averaging them. The number of these data points was 73 after cloudy pixels were removed.

The campaign took place during an algae spring bloom dominated by dinoflagellates and diatoms. The spatial variations of OSS were substantial within the measurement area. For a more detailed description of the data see [P3].

4.2.2 Remote sensing instruments and data

AISA spectrometer

AISA (Airborne spectrometer for applications) (Mäkisara et al., 1993) is an imaging spectrometer designed for airborne measurement although it can also be used for ground based campaigns. The main measurement characteristics of the AISA version used in this work are presented in Table 4.1. The LST currently owns an AISA. It can be installed on a Short SC7 Skyvan turboprop aircraft also owned and operated by LST.

Table 4.1. Measurement characteristics of AISA airborne spectrometer.

Type	Pushbroom CCD-matrix sensor
Number of channels	286
Channel wavelength range	450 – 900 nm
Channel bandwidth	1.6 – 9.4 nm (sum of 1 to 6 channels)
Number of pixels (across track)	384
Field of view	21°
Pixel size from 1000 m altitude	1 m
Pixel size after rectification	2 m

The total number of AISA channels is 286. However, the instrument is not able to store data from all channels when the measurement mode suitable for airborne remote sensing is used (the amount of data generated exceeds the capabilities of the data recorder). Instead, data from a smaller number of pre-selected channels are stored. In the campaigns of 1996-1998, the selected channels covered most of the total wavelength range (450 – 900 nm) although there were some gaps. Additionally, the channel configuration varied from campaign to campaign as the measurement requirements were defined more accurately. For example, the number of stored channels was 40 in campaigns conducted in August 1996 and May 1997 and 53 in August 1997 and August 1998. During the Vuosaari-campaign the number of channels was 32.

MODIS spectroradiometer

The most relevant technical details of MODIS are presented in Table 4.2 (Barnes et al., 1998). As Finnish lakes are fairly small and have an irregular shape the most interesting MODIS channels are the two that have the highest spatial resolution (250 m). These channels operate at 620-670 nm and 841-876 nm and are the only channels used in this work.

MERIS spectrometer

Table 4.2 shows the technical details of MERIS relevant to our study (Rast et al., 1999). The advantage of MERIS over MODIS is that MERIS has 15 channels with 300 m resolution while MODIS has 2 channels with 250 m resolution.

4.2.3 Data preprocessing

Preprocessing reduces the amount of data and makes it possible to numerically analyze (in our case with Matlab-software) the relationship between *in situ* and remote sensing data. The result of preprocessing is usually a matrix where remote sensing and *in situ* data are organized into columns and rows so that the remote sensing data from one geographical location matches the *in situ* data from that location.

For image (remotely sensed) data the preprocessing consists of radiometric and geometric corrections. Radiometric correction transforms the digital numbers recorded by the sensor into a physically meaningful unit, i.e. radiance or reflectance. Geometric correction rectifies the data so that they can be overlaid on a known coordinate system on ground. After the images are rectified it is possible to compare

the radiance value of a pixel with a corresponding ground truth values measured at an *in situ* station.

AISA images were corrected using GPS and inertial navigation system data and calibration measurements performed in a laboratory. AISA data were resampled to a pixel size of 2 m x 2 m. Additional data preprocessing consisted of deriving the average radiance of each AISA channel in a 100 m x 100 m square around each ground truth sampling point.

Figure 4.1 shows radiance spectra measured by the AISA airborne spectrometer at five ground truth data points in Lake Hiidenvesi. The values of water quality variables at these points are presented in Table 4.3. By comparing the spectra data with the *in situ* data some features can be observed:

- The peak at about 700 nm grows with the increasing concentration of chl *a*.
- The detected radiance increases with the concentration of total suspended matter in all parts of the spectrum.

These features are very similar to those visible in the simulated data (Figure 2.3).

Table 4.2. Characteristics of MODIS and MERIS instruments.

	MODIS	MERIS
Channels (wavelength coverage)	36 (405 nm –14 μm)	15 (400 - 1050 mm)
Resolution	250 m (2 channels) 500 m (5 channels) 1000 m (29 channels)	300 m
Swath width (km)	2330	1150
Time for global coverage (days)	1-2	3
Field of view	110°	68.5°
Scan type	Rotating double-sided mirror scan	Push-broom
Quantization (bits)	12	16
Design life time (years)	6	5
Satellite(s)	Terra (EOS AM) Aqua (EOS PM)	Envisat
Orbit	705 km, 10:30 a.m. descending node for Terra, 1:30 p.m. for Aqua, sun-synchronous, near-polar, circular	800 km (mean altitude), 10:00 a.m. descending node, sun-synchronous, near-polar, near-circular
Data available starting from	Terra: Feb. 24, 2000 Aqua: Jul. 4, 2002	Mar. 22, 2002

The spaceborne data (both MODIS and MERIS) were corrected using the calibration and geolocation data included in the data files. After these steps the radiance values for each *in situ* station were extracted.

For MODIS data also a 5 x 5 pixel area around station was extracted and the minimum radiance value found in that area was selected for further analysis. This method was applied in order to reduce the effect of rectification errors, noisy detectors, thin clouds and nearby land on the radiance values. After the extraction, clouds and land were detected and masked with the radiance thresholds using the two channels.

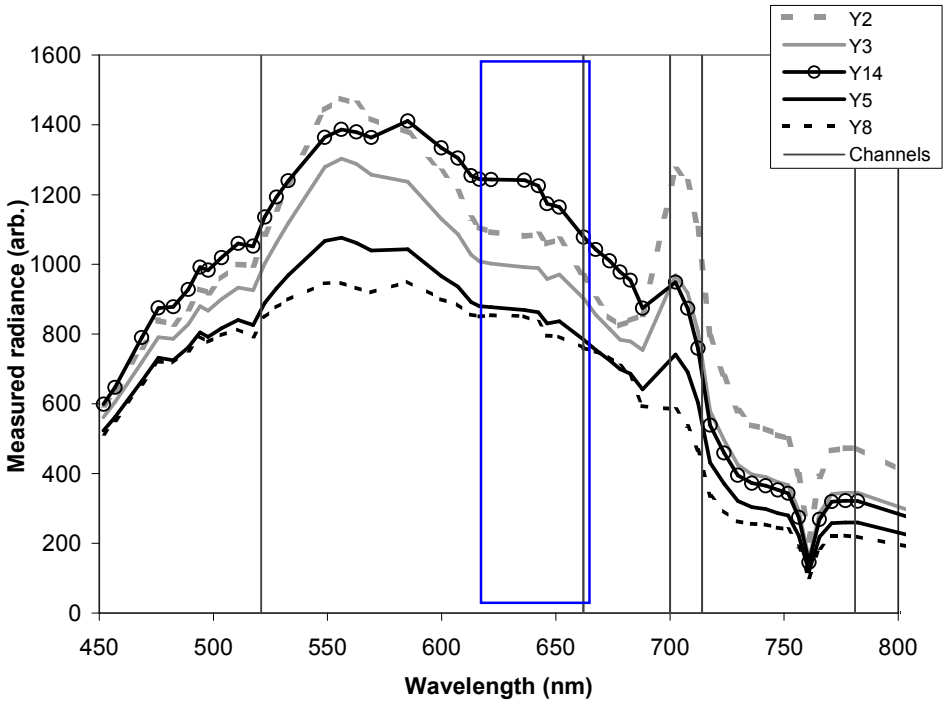


Figure 4.1. Sample spectra measured at Lake Hiidenvesi with AISA airborne spectrometer on August 11, 1998 [P5]. Y2-Y14 are in situ stations (the values of water quality variables at these station are presented in Table 4.3). The AISA-channels used in the retrieval algorithms (see Chapter 4.3) are shown as vertical lines. The downward absorption peak at about 760 nm is caused by atmospheric oxygen. The wavelength coverage of MODIS channel-1 (620-670 nm) is indicated with the blue rectangle. A comparison of the radiances in this region with the turbidity values presented in Table 4.3 shows the general trend of increasing radiance with the increasing turbidity (the only data point that does not behave in this way is Y14 where the concentration of chl a is low compared to turbidity).

Table 4.3. Values of water quality variables on selected in situ stations at Lake Hiidenvesi on August 11, 1998. $a_{CDOM}(400)$ was not analyzed at Lake Hiidenvesi in 1998. In August 1997 $a_{CDOM}(400)$ ranged between 5.3 to 6.0 m^{-1}. The concentration of total suspended solids (TSS, determined using GF/C glassfiber filter) is also shown although it was not used in the analysis [P5].

Data point	Secchi depth (m)	Turbidity (FNU)	Chl *a* (μg/l)	TSS (mg/l)
Y2	0.7	16	70	13
Y3	1	15	46	8.8
Y14	0.95	11	17	7.2
Y5	1.5	7.6	18	5.0
Y8	1.8	5.9	7.5	3.0

4.3 Empirical estimation of water quality parameters

4.3.1 Basic methodology and parameter selection

The values of water quality parameters are estimated from remote sensing data in papers [P1]-[P5]. The estimation of water quality parameters is in those papers based on empirical methods. Analytical methods were not used in [P1], [P2], [P4] and [P5] since the bio-optical model developed for Finnish lakes (Kallio et al. 2005, Kallio 2006) was not available when the papers were written. The model has been parameterized with chl *a* values that are below 72 μg/l. This is clearly below the values found in [P3].

The parameters estimated in papers [P1]-[P5] are chl *a*, TSS, $a_{CDOM}(400)$, Secchi depth and turbidity. The selection of the parameters in each paper was usually restricted by what data were available. For example, as shown in Figure 4.1 MODIS channel-1 does not cover the wavelengths where the peak caused by chl *a* appears. Figure 4.1 shows that MODIS channel-1 is suitable for TSS estimation. However, as seen in Table 1.4 the number of *in situ* TSS measurements is quite low when compared to turbidity. Thus, a logical choice in [P4] was to use turbidity in order to include as many data points as possible in the analysis.

4.3.2 Linear regression

Linear regression is perhaps the most often used empirical method in remote sensing of water quality and it was also selected for this work. The use of channel ratios in the regression algorithm has been found to be suitable for the retrieval of many water quality parameters (see Chapter 3 for references). The ratio of two channels reduces the effect of factors such as measurement geometry and atmosphere on the retrieval. The concentration of a water quality variable *C* can be expressed as:

$$C = a\left(\frac{R_1}{R_2}\right)^b + c, \tag{4.1}$$

where R_1 and R_2 are the remotely sensed reflectances and *a, b* and *c* are empirically determined regression parameters. If radiances are used instead of reflectances, the values of parameters *a, b* and *c* change. When reflectances (or radiances) used in Eq.

(4.1) are not atmospherically corrected the coefficients also contain information about the atmosphere between the target and the sensor. Thus, in practice the coefficients are valid only for the atmospheric parameters prevailing during the measurement.

AISA data were, for the first time, used for estimating water quality (the concentration of chl *a*) in [P1]. In [P5] they were also used for Secchi depth and turbidity. In order to find the model that explains the relationship between AISA observations and *in situ* chl *a* data in the best possible way, several simple linear regression algorithms (e.g. channel difference and channel ratio) were tried with all possible AISA channel combinations. All available channels from about 450 nm to about 800 nm were used (channel details are presented in Härmä et al., 2001). The best algorithm and channel combination was selected by comparing the coefficients of determination for each case. This procedure was performed in order to ensure that the best algorithms (wavelength combination) are found. For example, for chl *a* the two channels are usually in the wavelength range 660-705 nm. By using AISA data, which has many channels, it was possible to find the optimal ones.

Another possibility for determining the channels would be to use simulated spectra (see Figure 2.3) and to try to find the best channels based on the shape of the spectra. However, the bio-optical model was not available when [P1] and [P5] were written.

In later papers, [P2] and [P3], where AISA data are also employed, the channel combinations are based on earlier work. In [P3], empirical algorithms were also developed for MERIS data.

In some cases a single band is used instead of the channel ratio. This is the case for 250 m resolution MODIS data, which is used for estimating turbidity of lakes in Southern Finland [P4].

In [P1] the main objective of using multi-source data, e.g. HUTRAD microwave radiometer (Hallikainen et al., 1998) and wind speed data in addition to optical data, was to test if the effects of the water surface on the retrieval of chl *a* could be reduced. As can be seen in Figure 2.1 and in Eq. (2.1) the signal from the water surface affects the detected signal in the optical region. The magnitude of the effect is a function of the roughness of the surface and the geometry of the measurement. The roughness of the surface can in turn be expressed as a function of wind speed (Cox and Munk, 1954). The roughness also has an effect on the microwave emissivity of the surface and, therefore, can be detected with radiometers operating in the microwave region. Therefore, by modifying the retrieval algorithms of water quality parameters so that the surface effect is included as an independent variable (radiometer data, bidirectional scattering or wind speed) the accuracy of the algorithm may improve.

The empirical model that describes the relationship between *in situ* and remote sensing data can now be expressed as:

$$C = a\left(\frac{R_1}{R_2}\right)^b + cM + d, \tag{4.2}$$

where M are additional data and d is an empirical coefficient.

The retrieval algorithms are developed with the available data and the advantages of using non-optical data are analyzed by comparing the coefficients of determination for each case.

In addition to developing empirical algorithms that can be used to estimate the values of chl *a*, TSS and $a_{CDOM}(400)$ with AISA and MERIS data, the effects of changes in the atmosphere were estimated with a method similar to the one used for MODIS data in Chapter 4.4. The MERIS top of atmosphere (TOA) radiances were simulated using the measured TOA radiance values (i.e. the bio-optical model was not used here) and atmospheric parameters based on different atmospheric visibility values [P3]. The visibility ranges from 25 to 40 km on a typical summer day between 9 and 12 a.m. at the coastal weather stations. The closest weather station of Finnish Meteorological Institute observed visibility values ranging from 25 km to 35 km during the measurement campaign day. During the morning (9-12 a.m. local time), the visibility was 35 km. Hence, the visibility values selected for this case were 25, 35 and 40 km (in Chapter 4.4 the visibilities were 25, 40 and 60 km).

The simulated radiances were then used to derive new values for OSS using the algorithms developed with original data. The estimation error was computed from the difference of original and simulated OSS values. This analysis gives an estimate of the measurement error when algorithms developed for one atmospheric situation are used for another dataset without using *in situ* data for recalibrating the coefficients of the algorithm.

4.3.3 Training, testing and accuracy

A value of a water quality variable estimated with remote sensing methods is not very useful without some indication about its accuracy. The simplest way to estimate the accuracy is to compare the *in situ* values with the retrieved values with some basic statistical measure (e.g. correlation coefficient, coefficient of determination, bias and RMS (root mean square) difference).

The measurement error can be divided into two classes: (1) the root mean square error (RMSE) represents the random repeatability of the measurement, and (2) the bias represents the systematic error due to methodology or equipment. For a simple linear regression RMSE is defined as (Milton and Arnold 1995):

$$\mathrm{RMSE} = \sqrt{\frac{1}{N-2}\sum_{i=1}^{N}(X_{IS,i} - X_{RS,i})^2}, \tag{4.3}$$

where N is the total number of observations in the test data set, $X_{IS,i}$ is the observed *in situ* value and $X_{RS,i}$ is the value estimated with remote sensing.

Bias in turn is defined as

$$\mathrm{bias} = \frac{1}{N}\sum_{i=1}^{N}(X_{IS,i} - X_{RS,i}). \tag{4.4}$$

For operational use the bias is usually more serious since RMSE can be reduced by spatially averaging measurement pixels. That is not the case for the bias; a large bias can lead into serious errors in the data interpretation (e.g. the mean chl *a* concentration of a lake).

The retrieval accuracy of a remote sensing method can also be examined by dividing the available data into training and testing data sets. First, the retrieval coefficients are derived by using the training data set (includes both remote sensing and *in situ* data). Then these coefficients and the test remote sensing data set are used to derive estimates of water quality variables. Finally the estimated values are compared with

the testing *in situ* data. The measurement error of each data pair (remotely estimated value and the corresponding *in situ* value) is simply the difference between the values of estimated and *in situ* data. From the resulting vector of numbers a single statistical value of accuracy (e.g. RMSE) can be derived.

This method can also be used to test the use of the method for operational use, i.e. when concurrent *in situ* data are not always available. The available data can be divided into training and testing data set according to e.g. measurement date: the retrieval algorithms are trained with data from all but one of the measurement days and the data from that one day are used for testing. I.e. the procedure described below is used:

1. Use all data except data from one day for training the algorithm.
2. Use the data that was not included in training for testing.
3. Repeat steps 1 and 2 so that a different day is used for testing (and not included in training) each time until all days have been excluded from training and used in testing once (i.e., the number of loops is equal to the number of measurement days).
4. Compile a result vector from the test data and use that to compute error statistics.

Here this procedure is called 'daily testing' and it was used in [P1] and [P5]. The basic assumptions of the procedure are that the training data set includes all relevant circumstances and that the test data set does not deviate too much from them.

The data from different days are assumed to be independent from each other. Table 2 in [P1] shows the dates of the campaigns and in August 1996 there are three consecutive days. During these days some lake stations were observed (both *in situ* and remote sensing data) two or even three times. Thus, while it is possible that the training and testing data sets are not fully independent, the probability of this is low, since the other factors that affect the measurement (state of the atmosphere, solar zenith angle, and so on) have changed.

Another useful measure of accuracy is the confidence interval. For a single pixel the confidence interval M_s can be derived from (Milton and Arnold 1995):

$$M_s = t_{\alpha/2} \cdot S \cdot \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x}_{IS})^2}{S_{xx}}}\,, \tag{4.5}$$

where $t_{\alpha/2}$ is a coefficient based on T_∞ distribution (valid for normally distributed data), S is the RMSE for algorithm, n is the number of samples in the training data set, x is the remotely sensed turbidity value, $\bar{x}_{IS}$ is the mean value of the data used for the algorithm training, and S_{xx} is given by

$$S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x}_{IS})^2\,, \tag{4.6}$$

where x_i is the remotely observed turbidity for the training data point i.

For an estimated mean (e.g. the mean turbidity of a lake basin) the confidence interval M_m can be derived with:

$$M_m = t_{\alpha/2} \cdot S \cdot \sqrt{\frac{1}{n} + \frac{(x - \bar{x}_{IS})^2}{S_{xx}}} \,. \tag{4.7}$$

Finally, the standard error of a mean value of a remotely estimated parameter (SE_M) is defined as

$$SE_M = S \cdot \sqrt{\frac{1}{n} + \frac{(x - \bar{x}_{IS})^2}{S_{xx}}} \,. \tag{4.8}$$

The number of *in situ* data points collected during one campaign day is usually less than 20 as it takes time the perform the measurements and to travel from station to station. Remote sensing images on the other hand can have thousands or even millions of pixels. As shown in Eq. (4.3) the number of data points affects the statistical error of measurement so that a high number of points results in smaller error. Hence, remote sensing data has the potential to be more accurate than *in situ* data when mean values are considered.

This was tested using AISA data acquired over two lakes in Southern Finland (during the SALMON campaigns) [P2]. First, the AISA radiances are transformed into chl *a* concentrations using empirical algorithms developed with *in situ* data. The resulting thematic chl *a* maps (see Figure 4.2) are then used to evaluate the statistical accuracy of using remote sensing data in addition to *in situ* data and finally this result is compared to the case where *in situ* data are used alone.

4.3.4 Results

AISA

Dekker (1993) and Gitelson et al. (1993) noted that an algorithm using a ratio of two channels at about 680 to 710 nm (corresponding to chl *a* fluorescence and volume scattering from particulate matter (Gitelson, 1992; Schalles et al., 1998; Smith and Baker, 1978)) and at about 660 to 680 nm (chl *a* absorption) is suitable for monitoring chl *a*. This was also true with our data [P1]; the best results were obtained with an AISA channel ratio where the wavelength of the denominator is 673 nm and the wavelength of the numerator is 702 nm. Channels close to 673 nm and 702 nm also gave fairly good result. Therefore, the algorithm is not very susceptible to errors in the wavelength calibration of the instrument. A sample of a chl *a* map obtained with the ratio algorithm and airborne data is shown in Figure 4.2.

Applying atmospheric correction (HIRLAM numerical weather prediction model and MODTRAN simulations (Kallio et al., 2001)), and hence the use of estimated surface reflectance values instead of observed radiance, to the data did not improve RMSE significantly. However, it did reduce the bias error in seven of the 11 cases [P1]. This is important since the bias component is usually more harmful for an operational system. The atmospheric correction also reduced the standard deviation of the retrieval coefficients (*a* and *c* in Eq. (4.1)).

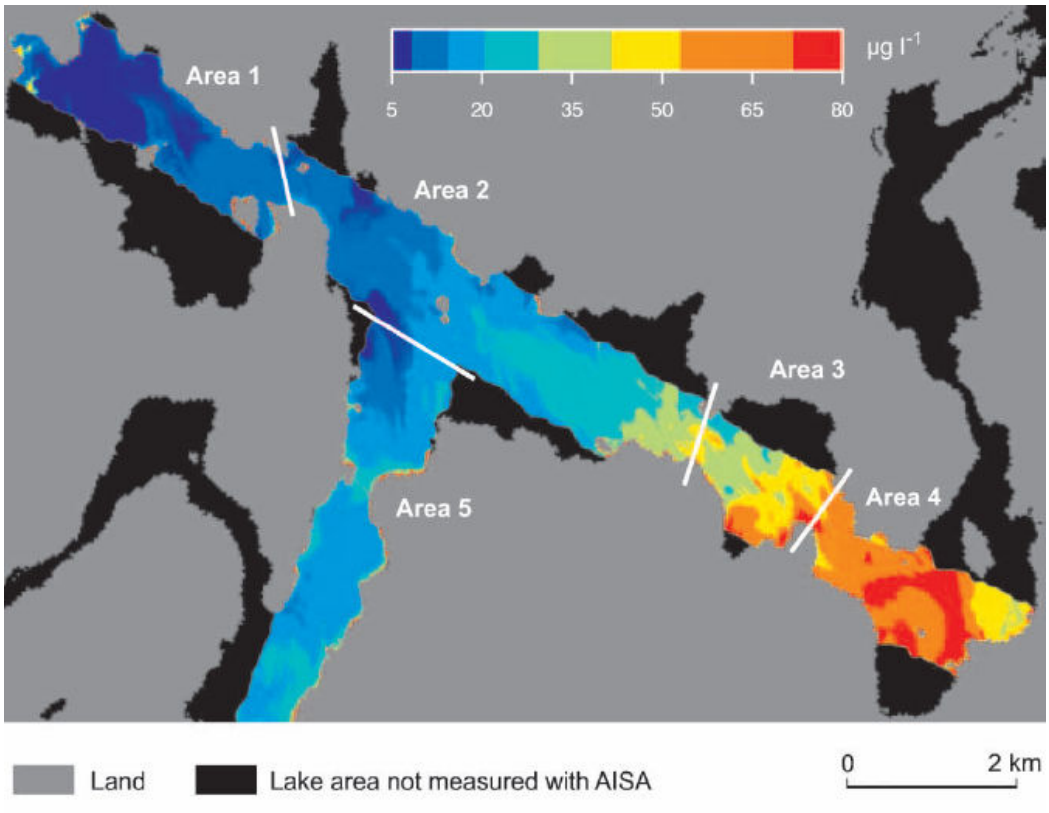


Figure 4.2. Chl a interpreted from AISA airborne spectrometer data at Lake Hiidenvesi on August 11, 1998. [P2]

Due to technical problems the availability of radiometer data varied between measurements. Therefore, an accuracy analysis with full data set was not possible and data from each measurement day had to be analyzed separately. The effect of the use of radiometer data on the accuracy of the retrieval was analyzed by comparing R^2 values with and without radiometer data. The improvement of R^2 was statistically significant in five of the eight cases. R^2 improved most in those cases where it was initially low.

Table 4.4 shows the RMSE and bias errors for the data acquired during 1997-1998 as various correction methods are employed. Wind speed data were available on four measurement dates (80 data points). The use of bidirectional correction data derived from wind speed data in the analysis reduced the bias and RMSE but only slightly in three of the four cases. The use of wind speed alone data did not improve the results.

Figure 4.3 shows one of the advantages channel ratio algorithms have over single channel algorithms. As can be seen the angular dependence (bidirectional effects) of channel ratios is much lower that the dependence of single channels. The difference is remarkable even for a relatively narrow field-of-view instrument like AISA. In the SALMON campaigns, the *in situ* stations were mostly close to the flight line of the aircraft. Thus, the effects of the different sensor angles to the results presented here are not large. Satellite sensors typically have a field-of-view in the order of 50 to 60 degrees and the angular effect is much larger for them.

The lake shown in Figure 4.2 can be divided into five sub-basins, which can be analyzed together or separately. The statistical characteristics the chl *a* distribution in these areas are shown in Table 4.5. During the SALMON campaigns, water samples

were collected from 15 stations within the lake. The lake also has three routine monitoring stations. The statistical characteristics of chl *a* at these stations are also shown in Table 4.5. Table 4.6 in turn shows the observed and standard errors for the mean chl *a* values of the lake derived for AISA based data and *in situ* data.

The chl *a* map in Figure 4.2 is a mosaic of two crossing images. As can be seen from the map, there is no border between the two images. This indicates that the chl-a interpretation algorithm is not sensitive to the change in measurement geometry (sun direction in relation to flight direction).

Table 4.4. Total RMSE (including bias) and bias error (in μg/l) for 1997-98 measurements with various correction methods [P1].

Date of test data (number of test data points)	No correction		Atmospheric correction		Bidirectional correction		Wind speed correction	
	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias
7 May 1997 (20)	5.9	5.0	6.9	6.1	5.2	4.4	15.3	14.6
11 Aug 1997 (38)	9.3	-7.1	11.4	-8.8	10.9	-9.0	9.5	-7.2
18 Aug 1997 (15)	3.5	2.4	3.7	1.5	3.3	-2.1	5.6	5.0
11 Aug 1998 (15)	7.0	2.8	8.5	5.3	6.3	-1.6	7.8	4.4
Mean	6.4	0.8	7.6	1.0	6.4	-2.1	9.6	4.2

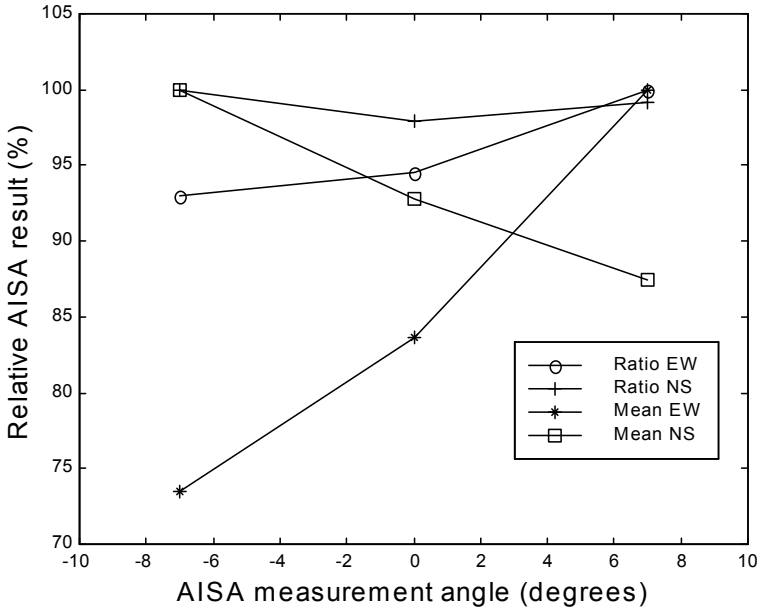


Figure 4.3. Relative channel ratio (702 nm/673 nm) and mean AISA-derived radiance (each measurement has been normalized to its maximum value in order to facilitate comparison) for flight directions East-West (EW) and North-South (NS) [P1]. The measurement took place in an area (Gulf of Finland) where water quality can be assumed to be homogenous within the swath of the sensor. Hence, the differences visible in the image can be assumed to be caused by measurement geometry.

Table 4.5. Statistical characteristics of the AISA based chl-a map of Lake Hiidenvesi on 11.8.1998. The last two columns indicate the statistical characteristics of the observed chl-a at 15 discrete in situ sampling stations and AISA-based chl-a at the three routine sampling stations. STD = standard deviation.

	Area 1	Area 2	Area 3	Area 4	Area 5	Areas 1-5	15 stations	Three stations
Number of pixels/stations	470 872	1 324 256	278 027	401 276	536 317	3 010 748	15	3
Area (km^2)	1.88	5.30	1.11	1.61	2.15	12.05	-	-
Mean chl-a ($\mu g\ l^{-1}$)	7.30	17.56	45.94	64.92	19.24	25.19	22.4	37.0
STD	2.05	7.15	11.59	10.58	2.80	19.6	19.7	27.6
Median ($\mu g\ l^{-1}$)	7	18	45	67	19	19	16	34
Min ($\mu g\ l^{-1}$)	2	3	18	37	11	2	6.2	11
Max ($\mu g\ l^{-1}$)	17	53	101	93	28	101	90	66

Table 4.6. Standard error and observed error of estimated mean chl a in Lake Hiidenvesi. Observed error was calculated as the difference between AISA based mean value and the mean of in situ values. The relative errors indicated in parentheses were obtained by dividing the absolute errors by the true mean chl a. The true mean chl a was assumed to be the AISA-based estimation (25.19 mg l^{-1}).

Data set	Mean chl *a* $\mu g\ l^{-1}$	Observed error $\mu g\ l^{-1}$	Standard error $\mu g\ l^{-1}$
AISA data	25.19	-	1.00 (4.0%)
15 stations	22.41	2.78 (11.0%)	5.08 (20.2%)
3 routine stations	37.00	11.8 (46.9%)	15.9 (63.3%)

MODIS

The data presented in Figure 4.1 and Table 4.3 indicate that MODIS channel-1 (620-670 nm) radiance ($L_{ch\text{-}1}$) increases with increasing turbidity. The best fit for $L_{ch\text{-}1}$ and *in situ* turbidity (T) data were obtained when [P4]:

$$\sqrt{T} \propto L_{ch\text{-}1}. \tag{4.9}$$

With that relationship the coefficient of determination (R^2) was 0.92. However, most of the data points used to derive the relationship have turbidity less than 6 FNU, and the three points that have high turbidity value dominate the regression. For about 84 % of the 1559 data points where turbidity was determined during summer 2000 the turbidity was less than 6 FNU; hence, the retrieval accuracy for low turbidity values is more important than the accuracy for high values. When only those points that have turbidity less than 6 FNU were used $R^2 = 0.69$ (data points closer than 50 m to shore removed).

Figure 4.4a shows the value of R^2 as a function of the minimum distance from the nearest shore (the data points used in the analysis are farther from the shore than the

value in the X-axis). The rise of R^2 as a function of distance is apparent. When the minimum distance is 450 m R^2 is about 0.9. Figure 4.4b shows the number of data points as a function of minimum distance. At 450 m 18 data points remain. Figure 4.5 shows a scatter plot of *in situ* turbidity and remotely sensed turbidity for these 18 points. The regression model derived from the 18 data points is (in FNU):

$$T = (0.52\, L_{ch\text{-}1} - 3.76)^2. \tag{4.10}$$

The RMSE of the data (derived with equation (4.3)) is 0.4 FNU (27 % of the mean turbidity). Figure 4.6 shows the 95 % confidence intervals for the turbidity model derived with the 18 data points. For a single pixel the interval is fairly large but for the mean of pixels the interval falls between 0.2 and 0.6 FNU for the most frequent turbidity range (0 to 6 FNU). This means that even when a single pixel might not represent the target in the best possible way the mean turbidity of a lake basin (several pixels) is generally correct.

Figure 4.7 shows the turbidity map of a region around Lake Päijänne as a visual example of the results remote sensing methods can produce.

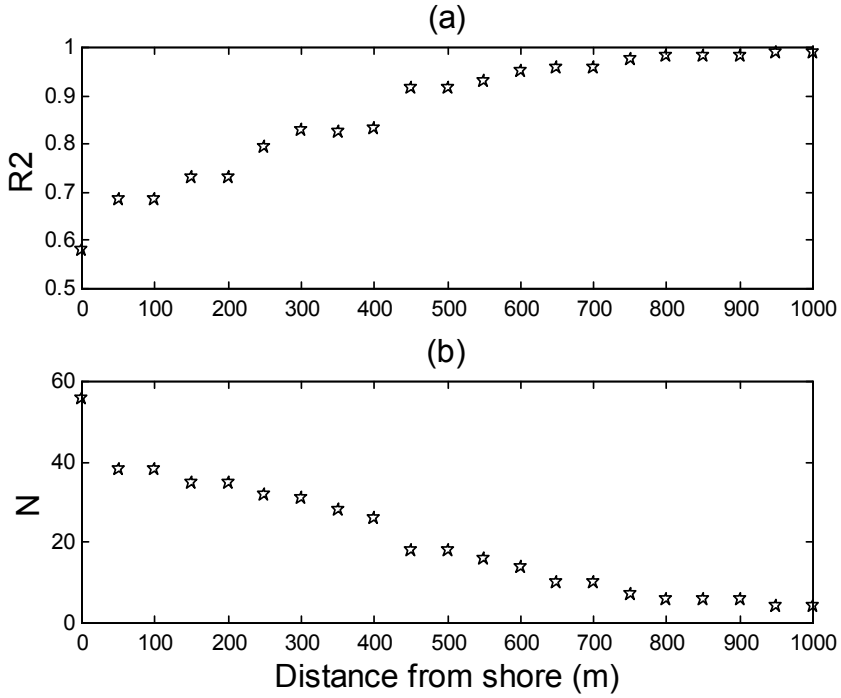


Figure 4.4 (a) The coefficient of determination (R^2) and (b) the number of data points (N) as a function of the minimum distance of the in situ data points from the nearest shore. The resolution of the land mask that was used in this analysis was 200 m. Eq. (4.10) was used as a relationship between radiance and turbidity. Points where the distance from shore is 0 m are in land according to the 200-m resolution land mask. Only points that have T < 6 FNU were used [P4].

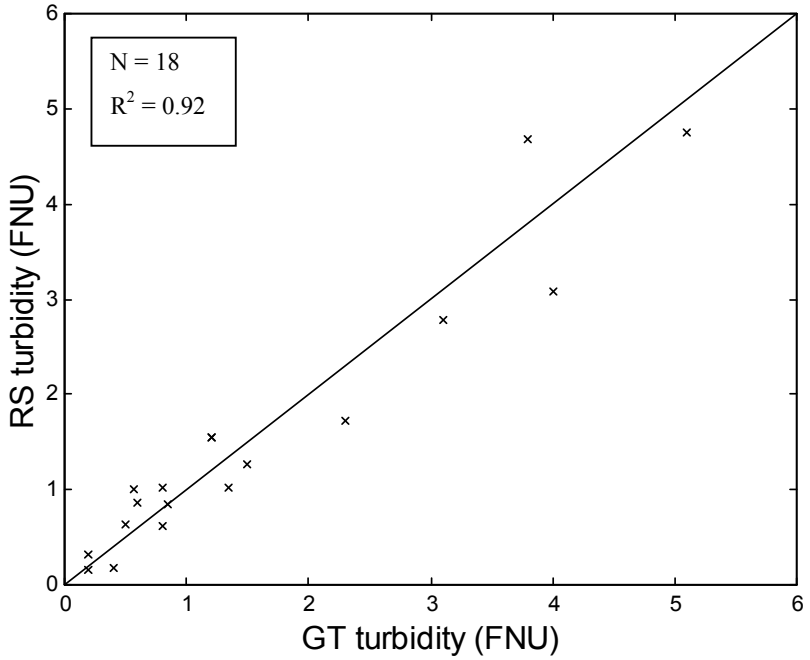


Figure 4.5. Ground truth (GT) turbidity vs. remotely sensed (RS) turbidity (derived with Eq. (4.10) by using MODIS data acquired on August 27, 2000 and ground truth data collected within ±3 days of the MODIS overpass). The minimum distance from shore is 450 m [P4].

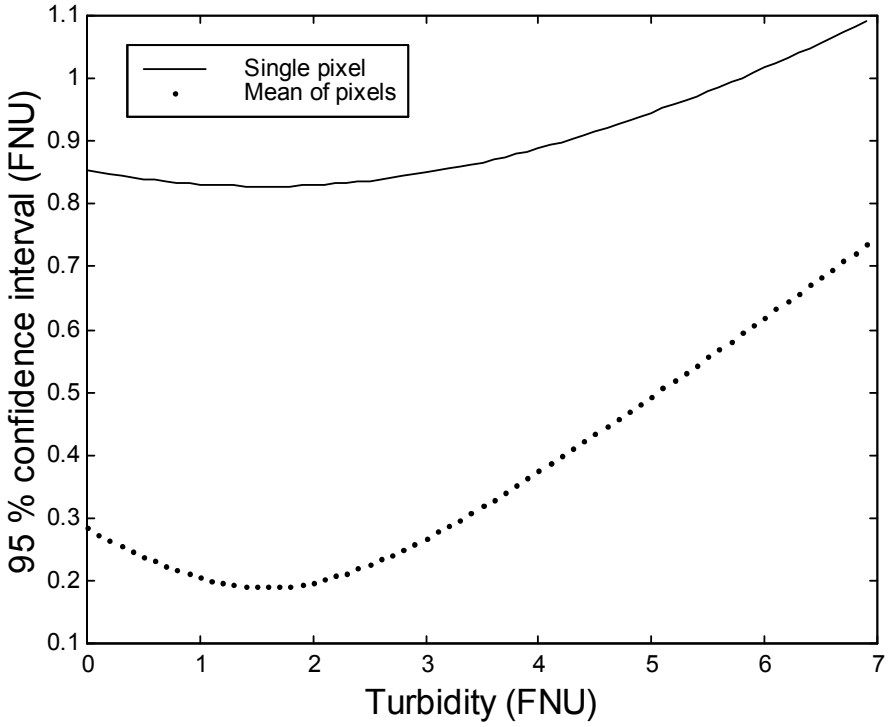


Figure 4.6. Confidence intervals (95 %) for turbidity regression model (N=18). The solid line is for single pixels (Eq. (4.5)). The dotted line is for the mean of pixel values (Eq. (4.7)) [P4].

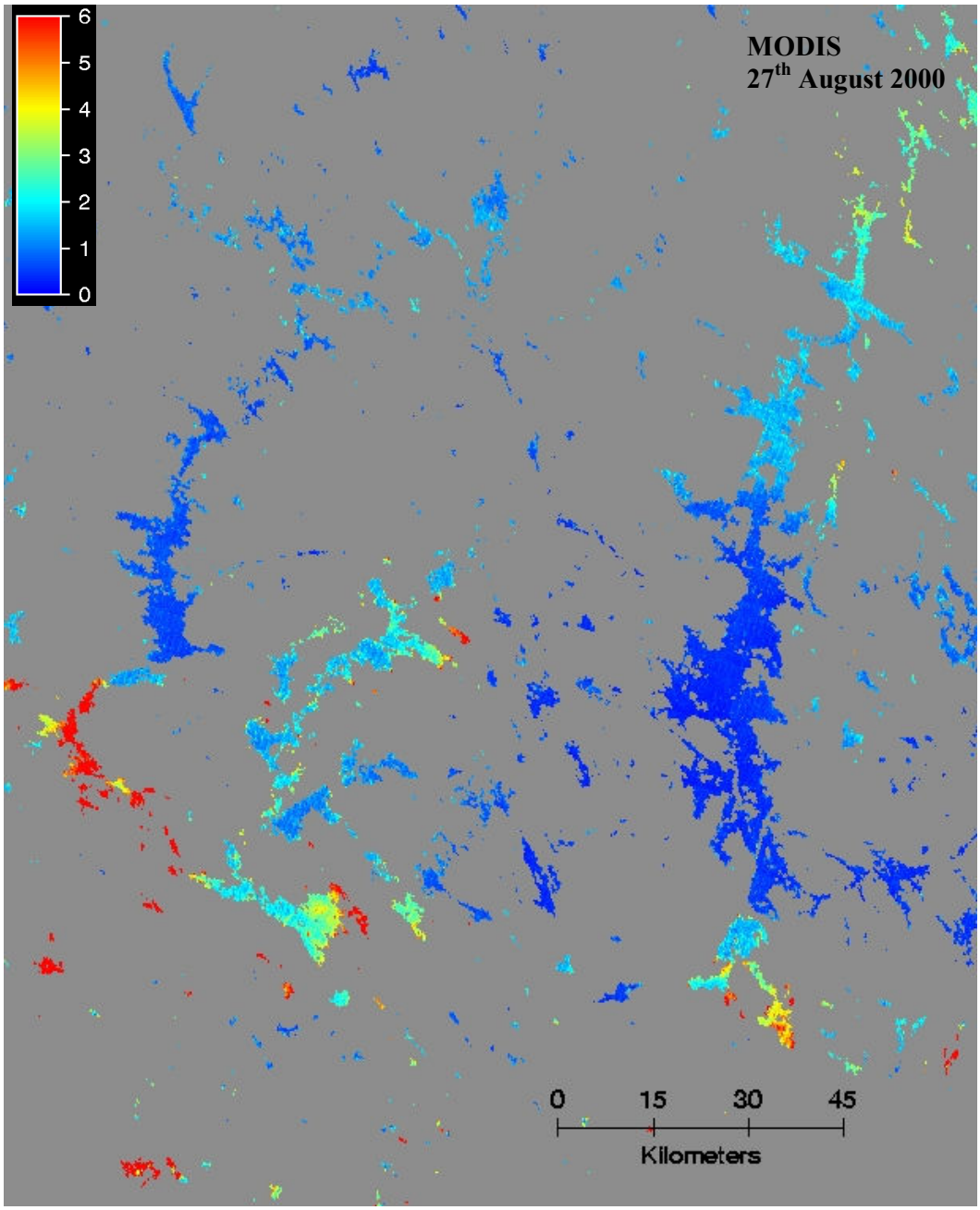


Figure 4.7. Turbidity map of Lake Päijänne region (the scale is in FNU units) as retrieved by using Eq. (4.10) and MODIS channel-1 data. Gray areas are land or clouds.

MERIS

Figure 4.8 shows a thematic map of chl *a* obtained with MERIS and AISA data on the coast of Finland [P3]. Figure 4.9 shows the values of TSS, chl *a* and $a_{CDOM}(400)$ estimated with MERIS and flow-through data as a function of the distance from the beginning of the boat transect. Both figures show the trend of decreasing chl *a* values as the distance from the shore increases. However, there are also areas where the chl *a* values increase and the remote sensing methods are able to detect those.

The algorithms for chl a, TSS and $a_{CDOM}(400)$ with MERIS and flow-through data are shown in Table 4.7. As can be seen the algorithms are based on channels where each OSS has an effect of the spectra as shown in Figure 2.3 (e.g. 709 nm for chl *a* and 490 nm for CDOM).The table also includes R^2 and RMSE values when all data points are used (n = 73) and R^2 values for a case where those MERIS pixels that contain less than 20 flow-through data points are excluded (n = 50). The results improve when only those pixels that have been sampled with several times by the flow-through device are used.

Figure 4.10 shows the simulated error of the chl *a* estimation when the atmospheric conditions are different from those that prevailed for the data used in development of the chl *a* algorithm. The error is well below 20% when the concentration of chl *a* is between 20 and 120 µg/l.

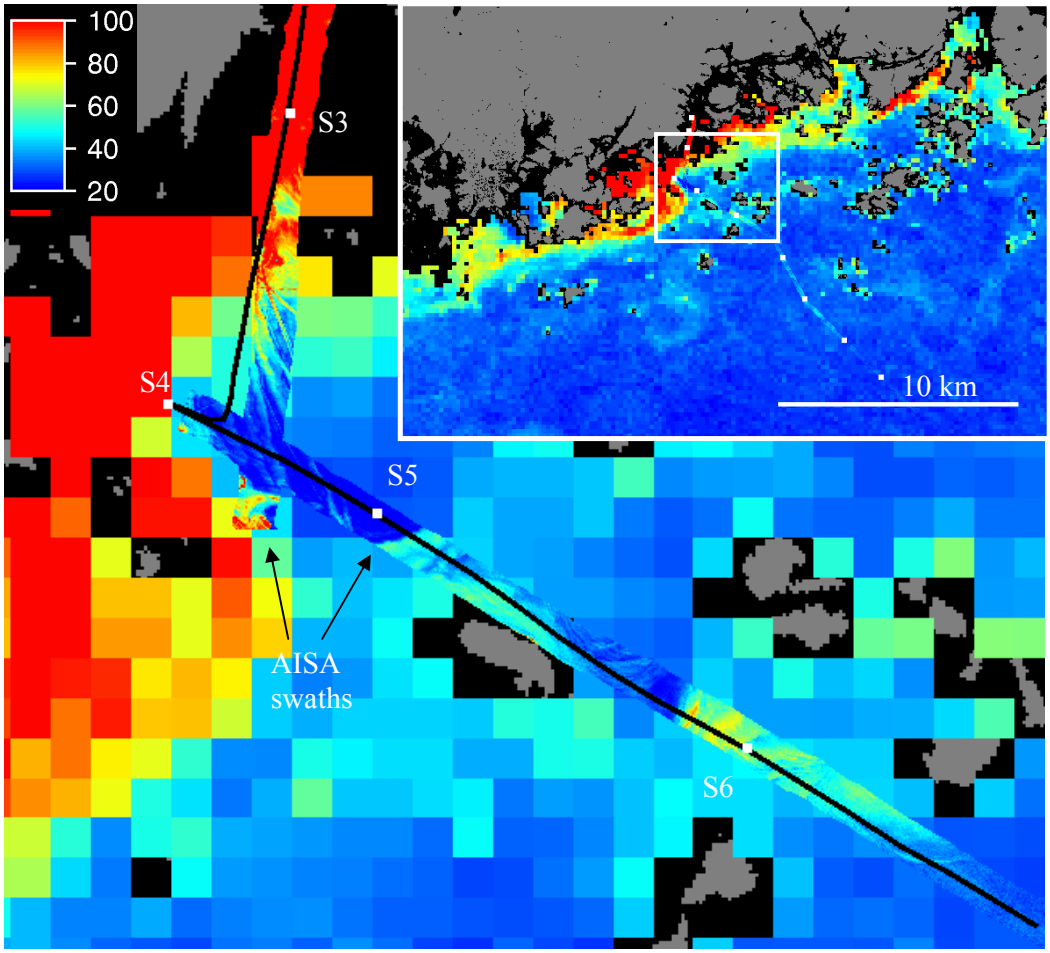


Figure 4.8. Chl a (in µg/l) with MERIS and AISA data. The concentrations over 100 µg/l are shown in red. S3-S6 are sampling stations. The small window shows an overview of the area, while the larger one shows the northern part of the airborne data where the most interesting features are present. The black curve shows the boat transect [P3].

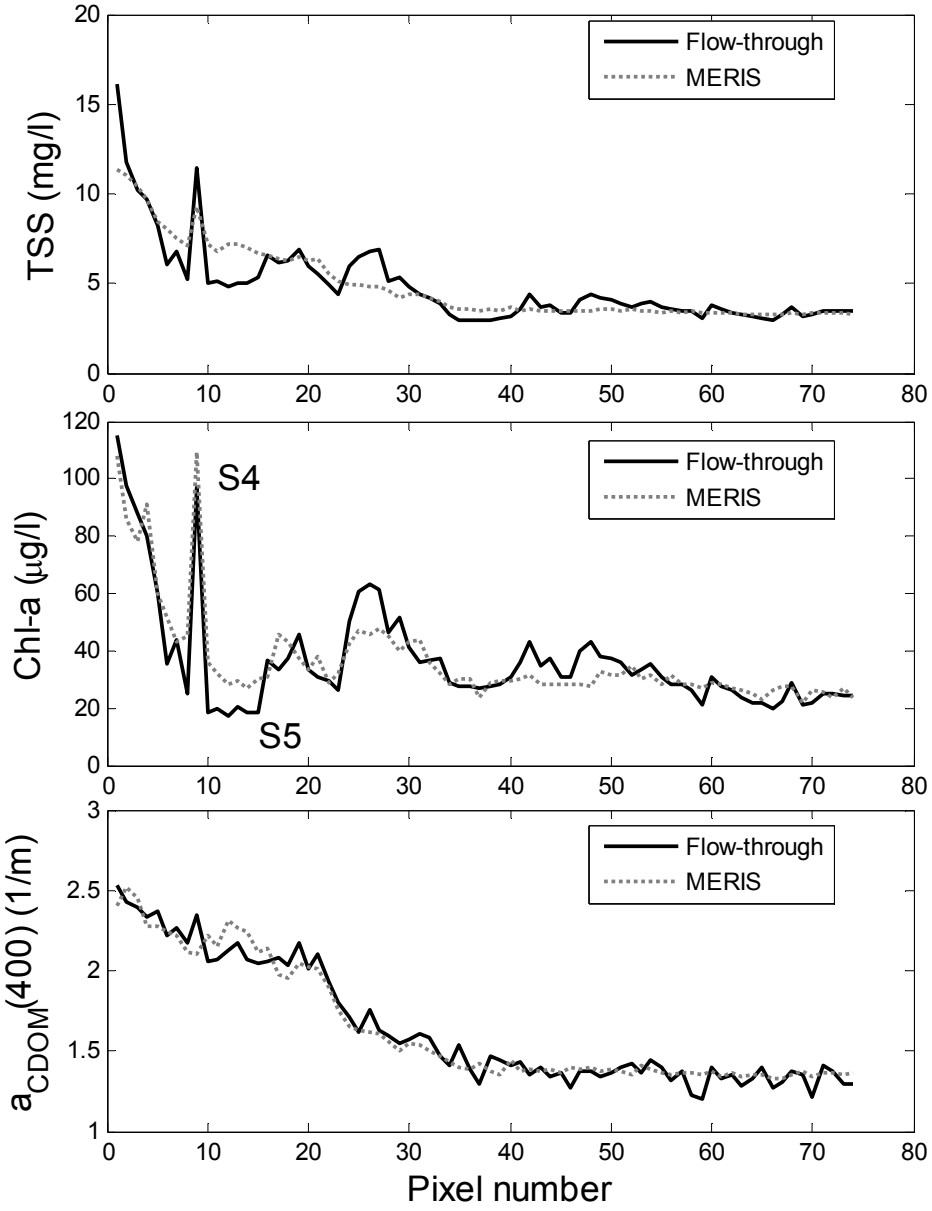


Figure 4.9. TSS, chl a and $a_{CDOM}(400)$ values with MERIS and flow-through data averaged within each MERIS pixel as a function of pixel number (roughly equal to distance).

Table 4.7. Retrieval algorithms, the coefficients of determination (R^2 in %) and the root mean square errors (RMSE) for MERIS and flow-through data during a coastal bloom event.

	Chl *a* (μg/l)	TSS (mg/l)	$a_{CDOM}(400)$ (1/m)
Algorithm	$275\frac{L_{709}}{L_{665}}$ - 189	$90.0\frac{L_{709}}{L_{560}+L_{665}}$ -19.6	$8.53\frac{L_{665}}{L_{490}}$ - 1.11
R^2 for all data (n = 73)	81.6	88.8	95.3
R^2 (%) with some pixels excluded* (n = 51)	86.7	92.3	96.0
RMSE (% from the mean value)	7.8 μg/l (22 %)	0.74 mg/l (16 %)	0.08 1/m (5 %)

* These results include only those MERIS pixels that contain 20 or more flow-through data points.

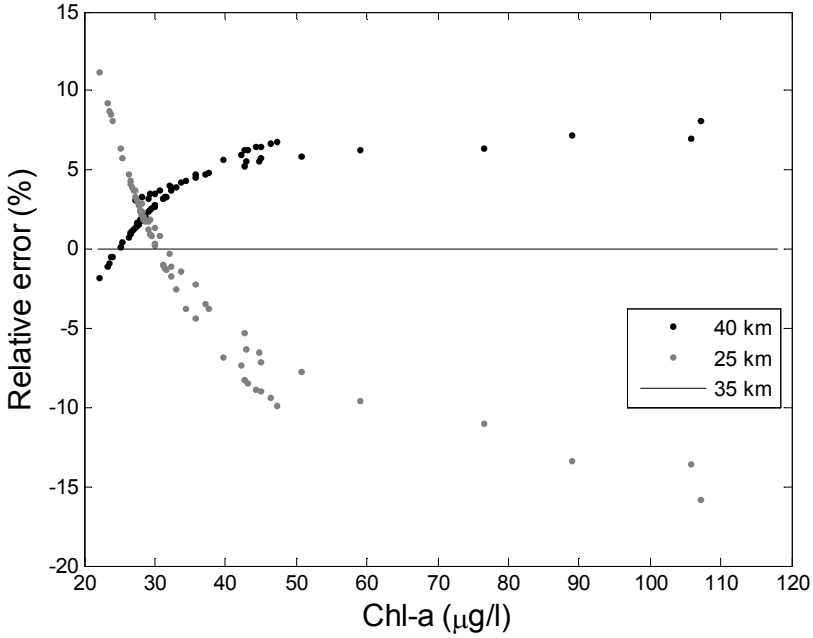


Figure 4.10. The relative (compared to the visibility = 35 km -case) errors of a fixed band-ratio algorithm for chl a under varying atmospheric conditions (visibility = 25 km or 40 km).

4.4 Lake classification

4.4.1 Basics of lake classification

Classification systems condense large amounts of data collected from water bodies into easily understandable information used by e.g. decision makers, authorities and the general public. Finding suitable variables and classification limits for a classification system is complicated. The variables and classification limits usually depend on the geographic location, the intended use of the water, data that are available and which organization is doing the classification. Classification can be

based on physico-chemical (e.g. chlorophyll *a*, total phosphorous, Secchi depth) and biological (e.g. species composition of phytoplankton, periphytic growth, macrophytes, and fish fauna) variables. An extensive review of different classification systems is presented by Premazzi and Chiaudani (1992).

Biological variables often describe the status of waters better than physico-chemical variables. However, the measurement of biological variables is expensive and time consuming. Hence, the operational classification of lakes in countries that have a high number of lakes (such as Finland) is usually based on physico-chemical measurements. The Finnish lake classification system (Heinonen and Herve, 1987; Vuoristo, 1998) uses about 20 water quality variables, of which the following three can be monitored with remote sensing instruments: chl *a*, Secchi depth and turbidity. The system also includes the occurrence of algal blooms, which can be visible in a remote-sensing image. However, the SALMON data set did not include algal bloom data and this variable was not included in the analysis.

By combining the Finnish system with the OECD Lake Classification Scheme (OECD, 1982; Premazzi and Chiaudani, 1992) a classification system suitable for remote sensing of Finnish lakes was developed [P5]. The limits for Secchi depth and turbidity were selected from the Finnish system, as the lakes in Finland are not as transparent as the lakes the OECD system was developed for. For chl *a* the Finnish system uses the mean concentration of the growing season, while the OECD system is based on the maximum concentration. As our data were from August when the chl *a* concentrations are usually at the maximum in Finland the OECD classification limits were selected for chl *a*. The variables and class limits of the combined system are presented in Table 4.8.

4.4.2 Methodology

In this analysis the classification was performed by first developing an empirical retrieval algorithm for each water quality variable separately (see chapter 4.3 and Table 6 in [P5]) and then using the limits presented in Table 4.8 to sort the data into water quality classes. The accuracy of the classification was assessed by forming classification matrixes for each variable and by comparing the percentages of correct classifications with and without the 'daily testing' procedure described below.

Lake classification was tested also with simulated MERIS data. The simulated MERIS values were derived by averaging AISA radiances located within the wavelength ranges of each MERIS band.

Table 4.8. Classification limits for airborne data [P5].

	Class 1	Class 2	Class 3	Class 4	Class 5
Secchi depth (m)	> 2.5	1 –2.5	< 1		
Turbidity (FNU*)	< 1.4	1.4 – 4.4	4.4 – 8.3	8.3 –19.6	> 19.6
(corresponding total phosphorous value in μg/l)	(< 12)	(12-30)	(30-50)	(50-100)	(> 100)
Chl *a* (μg/l)	< 2.5	2.5 - 8	8 - 25	25 –75	> 75

*FNU = Formazin Nephelometric Unit

Lake classification is also possible with a method that is much simpler than the one used with AISA data. The simpler method is based on the assumption that the quality class of a lake is directly proportional to the radiance value detected over the lake with for example the MODIS sensor [P6]. Instead of first estimating the values of the water quality parameters with remote sensing data the MODIS radiance values are used in the classification directly with empirical threshold values for each class.

The water quality characteristics are chosen based on the limits used in the AISA study presented above, except for chl *a.* The limits for chl *a* are based on the mean values of the growing season as the classification based on *in situ* data also uses those. As the characteristics used in the bio-optical model (chl a, TSS, and $a_{CDOM}(400)$) are slightly different from those used in the routine monitoring (chl *a*, water color, and turbidity) they had to be modified [P6]. The final limits are presented in Table 4.9.

The amounts of aerosols and gases present in the atmosphere vary both spatially and temporally. These variations can disturb the estimation of water quality characteristics when remote sensing methods are used. Hence, a bio-optical reflectance model (see Eq. (2.7) and Appendix A) (Kallio et al. 2005; Kallio 2006) combined with an atmospheric model presented in Pulliainen et al. (2000) is used to simulate the radiance values that water bodies with certain characteristics would reflect under various atmospheric conditions. The atmospheric model uses principal component analysis and MODTRAN simulations to reduce the variability of the atmosphere into one scalar variable, which is related to visibility. In this case, the parameters of the atmosphere ($T_{atm}(\lambda)$ and $L_a(\lambda)$, λ = 620-670 nm) are computed with the model using visibility values 25, 40 and 60 km.

4.4.3 Results with AISA and simulated MERIS data

The algorithms used for the classification are presented in Table 4.10. The data collected during May were not used for the classification as the retrieval algorithm for chl *a* is season specific. Therefore, the results are only valid for conditions that are similar to those observed in August (i.e. the fall period, see Chapter 1.1.1 for information about the yearly cycle of water quality). As an example of the classification of lakes with airborne data the results for chl *a* are shown in Table 4.11 (all data in training and testing) and in Figure 4.11. The difference between daily testing and training and testing with all data is noticeable only when the concentration of chl *a* is high. In other cases the two data sets match each other very well. Also, this did not affect the classification accuracy significantly as only one sample was misclassified by two classes. It represents Lake Tuusulanjärvi from which the data were not included in training the algorithm. This indicates that since there already are concurrent *in situ* and remote sensing data in archive those can be used for training purposes and *in situ* data may not be needed for every new campaign.

The results with the simulated MERIS data are very similar to those with AISA data. The classification accuracies are about the same as with the original algorithms using the full AISA spectrum information. This is not a surprise since the MERIS algorithms are almost the same as the AISA algorithms [P5]. The only differences are wider channels with MERIS and the exact center wavelengths used in the algorithms. A summary of the classification results is presented in Table 4.12.

Table 4.9. Classification limits used in the MODIS simulation [P6]. The value in the parentheses is the mean value of the class.

Class	Chl a (μg/l)	TSS (mg/l)	$a_{CDOM}400$ (m^{-1})
1	0-4 (2)	0-1.7 (0.85)	0-6.0 (3)
2	4-10 (6)	1.7-5.3 (3.5)	6.0-11.9 (8.9)
3	10-20 (15)	5.3-10.0 (7.7)	11.9-17.9 (14.9)
4	20-50 (35)	10.0-23.7 (16.9)	17.9-35.7 (26.8)

Table 4.10. Retrieval algorithms for AISA-data (SALMON campaign) [P5].

	Secchi depth	Turbidity	Chl *a* [2]
Algorithm [1]	$a_0 + a_1 \frac{L_{521} - L_{781}}{L_{700} - L_{781}}$	$a_0 + a_1 L_{714}$	$a_0 + a_1 \frac{L_{700} - L_{781}}{L_{662} - L_{781}}$
R^2 for all data (%)	92.6	85.4	93.7
Coefficients (all data in training)	$a_0 = -0.4298$ $a_1 = 1.0926$	$a_0 = -0.9203$ $a_1 = 0.0155$	$a_0 = -33.79$ $a_1 = 65.66$
Number of data points	102	99	80 (94 in testing)

[1] L_{xyz} is the detected radiance at a channel with a center wavelength of xyz nm.
[2] Data from Lake Tuusulanjärvi not used in training.

Table 4.11. Classification matrix for chl a (all data in training and testing) [P5]. RS denotes remote sensing and GT denotes ground truth.

RS GT	Class 1	Class 2	Class 3	Class 4	Class 5	Classification accuracy (%)
Class 1	**7**	2	0	0	0	78
Class 2	1	**2**	4	0	0	29
Class 3	0	4	**39**	3	0	85
Class 4	0	1	1	**27**	1	90
Class 5	0	0	0	1	**1**	50
All						**81**

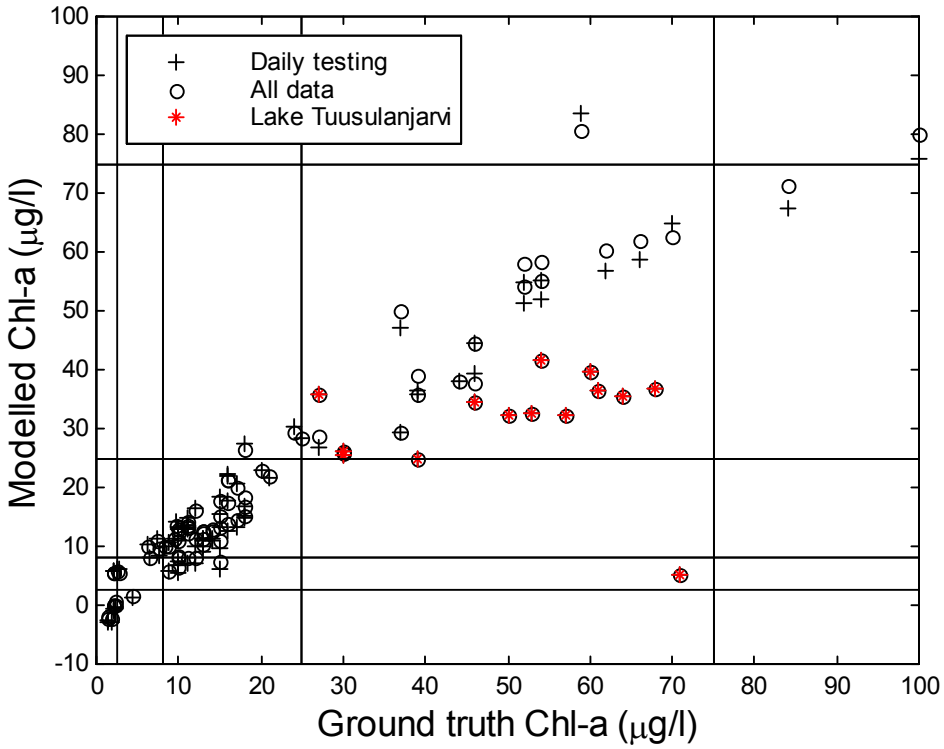


Figure 4.11. Regression model (see Table 4.10 for details) for chl a with AISA data when daily testing are used (+) and when all data have been used for training and testing (o). Vertical and horizontal lines represent quality class limits (five classes in total). Number of data points = 94. For all data R^2 = 93.7%, for daily testing (separate training and testing data sets) R^2 = 91.9% [P5].

Table 4.12. Classification accuracy for all cases (%) [P5].

Method	Secchi (3 classes)	Turbidity (5 classes)	Chl *a* (5 classes)
AISA (all data)	88	76	81
AISA (daily training)	90	79	78
MERIS	89	77	80

4.4.4 Results with MODIS data

The radiance values simulated with the combined bio-optical and atmospheric model are shown in Figure 4.12. As can be seen the amount of radiation reflected from the lakes increases linearly when the class number increases (and the water quality worsens). Figure 4.12 also shows the radiance values when the water quality parameters have the maximum and minimum values within each class, the effect that different atmospheric visibilities have on the radiance, and the combined effect of the class limits and the atmosphere. There is overlap in the radiance ranges of classes next to each other. However, when the class difference is more than one, the ranges do not overlap.

By using empirical radiance thresholds (see [P6]) MODIS pixels were classified into four classes. The results are shown in Figure 4.13 and in Table 4.13. The overall classification accuracy is 80.2 % and only 0.22 % of the data were misclassified by two classes. None of the data points were misclassified by three classes.

The sensitivity of the combined model on any single OSS is presented in Figure 4.14. The results are computed by allowing one OSS to change (the values are the mean class values Table 4.9) while the values of other OSS are kept constant. The analysis is repeated for each OSS and for the case where all of them are allowed to change. The largest impact on the radiance is caused by TSS. CDOM and chl *a* reduce the radiance as the class number increases. When the effects are combined the result is linear just like in Figure 4.12.

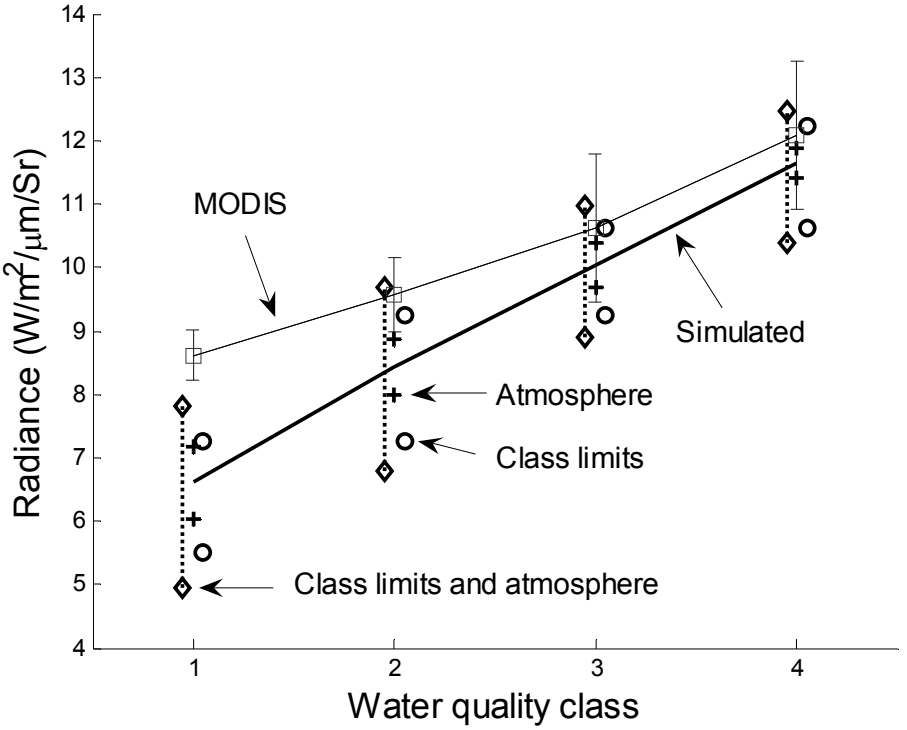
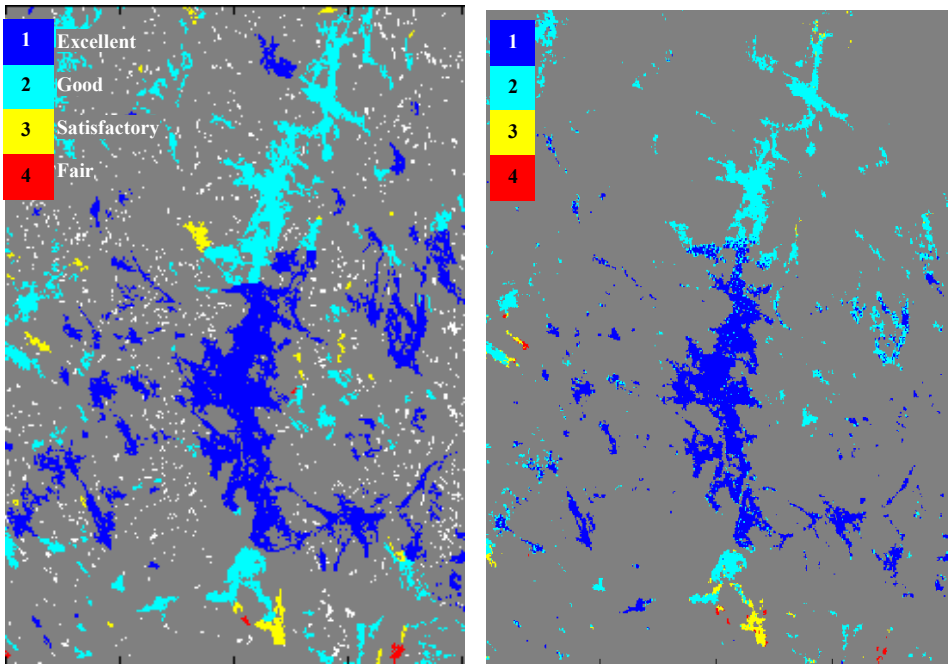


Figure 4.12. The simulated TOA-radiance (620-670 nm) using the mean, the minimum and the maximum values of water quality parameter within each class in Table 4.9 (visibility 40 km); the simulated radiance with atmospheric variation (visibility 25 and 60 km); the combined effect of class limits and the atmosphere; and the measured mean radiance (± standard deviation) of MODIS pixels as a function of the ground truth water quality class. [P6]

Table 4.13. Error matrix (number of pixels) for MODIS data. Total classification error: 19.8%; classification accuracy: 80.2%; error more than one class: 45 pixels (0.22 % of all pixels). The Totals column and row are the total number of pixels belonging to each class (the column is for the MODIS data and the row is for the ground truth data).

	Ground truth					
MODIS	Class 1	Class 2	Class 3	Class 4	Totals	Error %
Class 1	**9903**	954	24	0	10881	9.0
Class 2	2325	**6128**	276	7	8727	29.8
Class 3	1	364	**282**	6	653	56.8
Class 4	0	13	73	**44**	130	66.2
Totals	12229	7459	646	57	20391	
Error %	19.0	17.8	56.3	22.8		19.8



(a) (b)

Figure 4.13. (a) The quality of water in the Lake Päijänne region based on ground truth data from 1994-1997. White areas are lakes that are not classified. (b) Lake classification with MODIS data (August 27, 2000) using empirical threshold limits presented in [P6]. The colors of the classes are the same as in (a). The width of the area is 100 km.

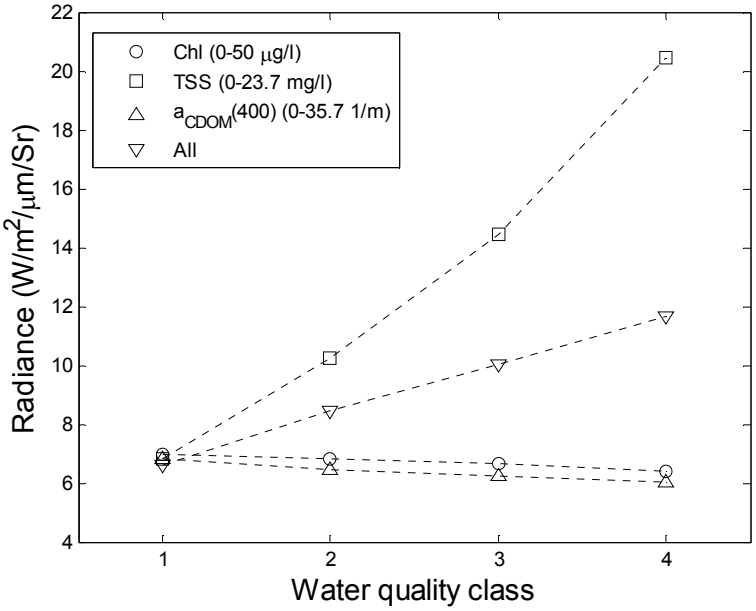


Figure 4.14. The sensitivity of simulated radiance (620-670 nm) on cases where one OSS (chl a, $a_{CDOM}(400)$ or TSS) is allowed to change while others are kept constant. The data labeled "all" is for a case where all are allowed to change according to values in Table 4.9. (Note: these results were not included in [P6] due to paper length limitations.)

5 Discussion

In this chapter the results presented above are compared with the work performed by other researchers and the implications of the research are discussed.

5.1 Estimation of water quality parameters

The airborne and spaceborne remote sensing methods employed in this work were successful at retrieving information about water quality from lakes and coastal areas of Finland. In the case of lakes, the airborne water quality monitoring system based on the AISA airborne spectrometer was able to retrieve the values of three water quality parameters (chl *a*, Secchi depth and turbidity) with good accuracy (R^2 above 0.85) despite different measurement configurations and lake types, and a large time span (measurements conducted 1996-1998). In the coastal case (a single day measurement during a spring bloom) airborne data were used for mapping chl *a*, TSS and $a_{CDOM}(400)$ with good accuracy (R^2 above 0.90, see Table 7 in [P3]).

The results with airborne data also showed that the use of multi-source data improves the accuracy of retrieving chl *a*, one of the most important water quality variables. However, since it is quite rare for a satellite to have both optical and microwave radiometers on board the possibilities for further studies of this method will be quite limited. Also, the limited spatial resolution of satellite microwave radiometers presents a problem.

The current operative lake monitoring system used in Finland is based on measurements at fixed stations. These stations (or in some cases just one station per lake) may not always present the actual condition of a lake in the best possible way. Perhaps the worst flaw of the current monitoring system is that the spatial resolution is limited. With remote sensing instruments it is possible to see how the values of water quality variables are distributed spatially and, thus, get information on the status of the whole lake. Information on the relative spatial variations of water quality variables is also interesting even though the absolute accuracy of a single pixel might not be as good as with laboratory techniques.

The statistical analysis with chl *a* data shows that the use of remote sensing methods together with *in situ* data improves the accuracy of the determination of the mean chl *a* of a lake by a factor of five (the error decreased from 20.2 % to 4.0 %). Also, the minimum and maximum values of chl *a* found within a lake can be estimated more accurately with remote sensing data than by using *in situ* data alone.

The retrieval accuracies (values of R^2; see Table 4.10) presented in this thesis for airborne data are at par or better than the results obtained by other authors who have used airborne instruments in their studies (see Table 3.1), especially when the extent of the data set (number of lakes and data points, and variability range of *in situ* data) used here is considered. For example, the R^2 of chl *a* obtained by Thiemann and Kaufmann (2000b), Östlund et al. (2001), George and Malthus (2001) ranges from 0.89 to 0.97 with 15 - 30 data points while our result was 0.94 with 80 data points. Also the variability of the data is large in our analysis. The data set used here covers all trophic states the lakes in Finland have (chl *a* varies from 1 to 100 μg/l), while e.g. in George and Malthus (2001) the chl *a* values have a mean of 5 μg/l and a maximum

of 10 μg/l and the data range in Östlund et al. (2001) was 2.9 - 48.1 μg/l. In Thiemann and Kaufmann (2000b) the data range was the same as in our case (1 to 100 μg/l).

The results with spaceborne data show that even with a simple method MODIS data are useful for monitoring turbidity in Finnish lakes. The usability of MODIS 250 m data for water quality estimation has also been confirmed in e.g. Miller and McKee (2004).

With the MERIS instrument the values of chl *a*, TSS and $a_{CDOM}(400)$ were successfully mapped during a spring bloom, although not as accurately as with airborne data. Spatial averaging due to larger pixel size removes part of the variation visible in the airborne data.

The comparison of the results obtained here for lakes with the results of other authors (Table 3.1) using spaceborne data (mostly Landsat-TM) is difficult, as the data and the methods are so different. However, some qualitative comparisons can be made. Here the number of data pairs for the MODIS turbidity analysis is 18 while in Strömbeck et al. (2003), Giardino et al. (2001), Östlund et al. (2001) and Thiemann and Kaufmann (2000a) it was twelve, four, six and eleven, respectively. Here the turbidity value of a large portion of the lakes in Southern Finland is mapped while Strömbeck et al. (2003), Giardino et al. (2001) and Östlund et al. (2001) only mapped one lake. Thiemann and Kaufmann (2000a) mapped about 60 lakes by using data collected from five lakes. Therefore, the method presented in this thesis is geographically more extensive.

For coastal areas (Table 3.2) the most popular instrument used in other studies, especially for chl *a*, has been SeaWiFS. The most popular chl *a* algorithm is the blue-green ratio (i.e. the Case 1 algorithm). The R^2 values are comparable to the results presented in this thesis (roughly 0.8 for chl *a*), however, the variability range of the data in this work is much larger than what is presented in other sources. Here the chl *a* range is from 20 to 120 μg/l, while in most other publications the maximum value is below 10 μg/l. Thus, the spring bloom case presented here is clearly exceptional when compared to data from other coastal areas. Successful monitoring of blooms like it requires algorithms calibrated for Baltic Sea conditions as the global methods tend to fail in this region. For example, the standard level 2 MERIS product produced by ESA includes values for chl *a* concentration. The neural network, which is used in the processing, has been trained to resolve OSS properties for a certain range of OSS values. For chl *a*, the valid range is only from 0 to 25 μg/l. This is clearly insufficient for the Gulf of Finland, where high OSS values are regularly observed during the spring bloom season. Thus, one purpose of the analysis described here is to explore the possibilities of OSS detection beyond the range of applicability of the current standard ESA processing methods.

5.2 Classification

The accuracy of the classification of lakes according to the airborne retrieval was also good (close to or over 80 % depending on the variable and the method). The accuracy remained high for all three variables even when the retrieval algorithms were tested using temporally separate test data sets neglecting the nearly simultaneous reference data in training the algorithms. This suggests that the current database should be large enough for operational airborne monitoring of the lakes used in the survey. However, the results are only valid for conditions that are similar to those observed in August.

The limited data from May indicate that the algorithms may vary seasonally due to e.g. different phytoplankton species composition (Kallio et al., 2001). Therefore the algorithms presented here should be first thoroughly tested before applying them to other seasons. The system developed here should also work well for lakes that are similar to the ones used here. For other lake types more concurrent *in situ* and remote sensing data are needed. However, Finnish lake types are fairly well covered in the present data set.

The use of MODIS and simulated Envisat MERIS data also yielded good results for lake classification with classification accuracy of close to or above 80 % for both instruments. The classification analysis with MODIS data was based only on one day; hence, it was not possible to test the daily testing procedure that was done with airborne data. However, the simulated radiances obtained with a combined bio-optical and atmospheric model were close to the MODIS values. This supports the validity the empirical method.

The accuracy of a classification system also depends on the number of classes the system uses. In this work the number of classes is five or less and part of the success may be attributed to that. However, in most cases no information at all is available from smaller lakes so even a coarse classification is useful. Furthermore, the experts who generated the operative classification systems discussed here have only used at most 5 classes (Heinonen and Herve, 1987; Vuoristo, 1998).

5.3 Effects of instrument characteristics

Three quite different remote sensing instruments were used in this work. Thus, it is possible to draw some conclusions on how instrument characteristics affect the estimation of water quality in Finland.

As shown in Figure 4.2 and Figure 4.8 the spatial variations within a lake or at a coastal area can be quite large even across a short distance. Instruments with different pixel sizes see these variations differently. If the pixel size is large some of the variation is lost due to averaging. Airborne instruments can offer the best resolution and thus see the highest and the lowest values in the data well. However, they are expensive to operate and can cover only a limited area.

Large spatial variations, when combined to large pixel size, can also make it difficult to compare remote sensing data with *in situ* data. This is true especially during a surface algal bloom (Kutser 2004). If the sample is collected from a local minimum or maximum and the remote sensing data sees the average over a larger area the data pair will be an outlier in a scatter plot. This can explain some of the error in the results presented here with satellite data. Unfortunately, it was not possible to quantify the amount of error caused by this. *In situ* sampling based on flow-through measurements is one (partial) solution to this problem.

Landsat-TM and other high-resolution satellite instruments have better coverage than airborne instruments and have been quite popular in lake remote sensing. However, data from high resolution instruments is also often costly (compared to MODIS data that can be downloaded free of charge), and the instruments have long revisit times (16 days for Landsat-TM compared to one day with MODIS), and low radiometric resolutions (8 bits for Landsat-TM compared to 12 bits for MODIS).

Another problem related to spatial resolution is that Finnish lakes are typically small and irregular in shape and may include small islands. The coastline in turn is filled with bays, peninsulas and islands. The radiation reflected from the shore and the vegetation near the shore is usually stronger than the radiation from water. Therefore, if even a small portion of a pixel is covered by land the retrieval of water quality variables may not be possible. Thus, as the size of the pixel grows the number of lakes that can be monitored diminishes.

The 300-m nadir resolution of MERIS and 250-m nadir resolution of MODIS are good enough for large and medium sized lakes. However, small ones cannot be monitored with them at all. To balance this drawback MERIS and MODIS have better temporal coverage (frequent overpasses) and in case of MERIS a good channel combination. The 250-m MODIS channels can only be used for monitoring TSS, turbidity and the general quality class of water. The channels available in MERIS can be used for estimating all three main OSS (chl *a*, TSS and CDOM). In Finnish lakes, the amount of CDOM can be high and thus it is important to monitor.

MODIS does have other advantages. One is that there currently are two MODIS instruments in orbit, which increases the amount of available data. This is very important in Finland due to the high probability of cloud cover. The other is that MODIS data can be freely downloaded a few days after the satellite overpass while MERIS data has to be ordered and they are delivered through CD-ROMs, which sometimes takes weeks. Furthermore, the Finnish Meteorological Institute (FMI) has a receiving station in Sodankylä (Northern Finland). The station can provide MODIS data to users within a few hours of the overpass.

The radiometric resolutions of the instruments used here were adequate. At least, the quantization of the data is not visible in the thematic maps.

As seen above, each instrument has its advantages and disadvantages. Thus, it is difficult to determine which of the current instrument is ”the best” for water quality remote sensing. It is usually better to choose the instrument based on what the objective of the measurement is (i.e. what water quality parameters are used, what is the size and shape of the water body, how often measurements are needed and so on).

If a new satellite instrument is designed specifically for lakes similar to Finnish ones, it should have spectral characteristics (channel positions and widths, and radiometric resolution) similar to those of MERIS. In order to increase the number of lakes that can be monitored the spatial resolution should be better than that of MERIS. Good temporal coverage is also essential due to the high probability of cloud cover. The number of images acquired by the instrument from any area should be at least two or three per week. It is possible that these requirements cannot be achieved with a single sensor and a constellation is needed instead.

5.4 Limitations of remote sensing

The difficulty of accurately performed atmospheric correction is usually seen as the largest obstacle on the path to successful multi-temporal water quality estimation with satellite sensors. For single date cases (such as the ones presented in [P4], [P6] and [P3]) the lack of atmospheric correction is not fatal if concurrent *in situ* data are available. In that case, the empirical regression coefficients computed using the *in situ* data include the effects of the atmosphere.

When two or more satellite images with different measurement geometries and atmospheric constituents are used in the estimation without concurrent in *situ data* for all images some kind of atmospheric correction is needed. It can be based on simple dark pixel subtraction as in Miller and McKee (2004), or on a more elaborate system of combined bio-optical and atmospheric model as outlined in Vepsäläinen et al. (2005). However, even when atmospheric correction is not performed the use of channel ratio algorithms reduces the errors caused by the atmosphere as shown in [P3].

The remote sensing of water quality with optical instruments has one more drawback: only those substances that scatter or absorb optical radiation can be monitored. Hence, some important parameters presented in Table 1.1, such as heavy metals and other poisonous substances, cannot be detected with optical instruments.

5.5 Final conclusions

Remote sensing is the only method that can provide spatially and temporally extensive data with reasonable accuracy. Hence, despite the limitations discussed above it is a valuable tool for water quality estimation. The results presented here improve the knowledge about remote sensing methods and their applicability for water quality estimation in Finnish conditions. For example, the classification of a selection of lakes as large as here and with as many water quality parameters as here has not been attempted before. Also, the use of the 250-m channels of MODIS (primarily used for land, cloud and aerosol boundaries) for lake monitoring is a novel idea.

Since remote sensing cannot alone provide all information required for water monitoring *in situ* monitoring will be necessary also in the future. An advanced water monitoring system will utilize data from both sources: *in situ* observations will provide accurate data on all necessary water quality parameters from a limited number of stations while remote sensing data are used to expand these measurements spatially and temporally. This will allow a greater portion of lakes and coastal areas to be monitored more frequently than with the present system.

The methods presented here are applicable to other regions (similar band ratio algorithms have been used for, e.g. chl *a* estimation in other European countries). However, since empirical methods are used the algorithms must be calibrated with new *in situ* measurements.

6 Future research

The use of remote sensing for estimating water quality in Finland is still a relative young research area. At the Laboratory of Space Technology and Finnish Environment Institute it started with the SALMON-project in 1996. Thus, the research on this topic will continue. The emphasis will be on the following:

- Improvement of MERIS and MODIS algorithms using multi-temporal data. The results presented here with spaceborne data are for single day cases and while those are valuable the aim is to provide the end users with reliable time series of data. The problems of atmospheric correction and automatic thin cloud detection (including jet aircraft contrails) must be solved so that the system is less dependent on *in situ* measurements.
- Use of MERIS data for lakes. The 250 MODIS data cannot be used for reliably estimating chl *a* and CDOM. Due to their small size and irregular shape the resolution of 1 km MODIS channels is too coarse for most Finnish lakes. Hence, the 300 m MERIS data will be used for estimating chl *a* and CDOM in lakes. This work has already started in the Tekes funded CATCH_LAKE-project.
- Improvement of the operational usability of remote sensing by assimilation and modeling. Remote sensing observations will be used to calibrate and optimize dynamic lake water quality models, which often lack the information on the spatial distribution of water quality variables.

7 Summary of appended papers

Paper 1

Chlorophyll-a (chl *a*) concentration of lake water can be measured with airborne (or spaceborne) optical remote sensing instruments by using empirical algorithms based on a channel ratio where the radiance of a channel at 680 to 710 nm is divided by the radiance of a channel at 660 to 680 nm. Airborne Imaging Spectrometer for Applications (AISA) was used in four lake water quality measurement campaigns (eight measurement days) in Southern Finland during 1996-1998 along with other airborne instruments and extensive *in situ* data collection. The RMSE obtained here for chl *a* estimation with empirical algorithms and 122 measurement points was 8.9 μg/l (all points used for training and testing).

As empirical algorithms are employed for chl *a* retrieval from remote sensing data, temporally varying factors such as surface reflection and atmospheric effects degrade the estimation accuracy. This paper analyzes the quantitative accuracy of empirical chl *a* retrieval algorithms available as methods to correct temporal disturbances are either included or excluded. The aim was to evaluate the usability of empirical chl *a* retrieval algorithms in cases when no concurrent reference *in situ* data are available. Four methods to reduce the effects of temporal variations are investigated. The methods are: (1) Atmospheric correction, (2) synchronous microwave radiometer data, (3) wind speed data, and (4) bidirectional scattering model based on wind speed and sun angle data.

The effects of different correction methods are analyzed by using single-date test data and multi-date training data sets. The results show that the use of a bidirectional scattering model and atmospheric correction reduces the bias component of the measurement error. Radiometer data also appear to improve the accuracy. However, if concurrent *in situ* reference data are not available, the retrieval algorithms and correction methods should be improved for reducing the bias error.

Paper 2

The spatial distribution of the sum of chlorophyll a and phaeophytin a concentrations (chl *a*) under light wind (0–2 m/s) conditions was studied in two lakes with an AISA airborne imaging spectrometer. Chl *a* was interpreted from AISA radiance data using an algorithm based on the near-infrared (700–710 nm) to red (660–665 nm) ratio. The results of Lake Lohjanjärvi demonstrate that the use of one monitoring station can result in over- or underestimation by 29–34% of the overall chl *a* compared with an AISA-based estimation. In Lake Hiidenvesi, the AISA-based estimation for the mean chl *a* with 95% confidence limits was 25.19±2.18 μg/l. The use of AISA data together with chl *a* measured at 15 *in situ* sampling stations decreased the relative standard error of the mean chl *a* estimation from 20.2% to 4.0% compared with the use of 15 discrete samples only. The relative standard error of the mean chl *a* using concentrations at the three routine monitoring stations was 15.9 μg/l (63.1%). The minimum and maximum chl *a* in Lake Hiidenvesi were 2 and 101 μg/l, 6 and 70 μg/l and 11 and 66 μg/l, estimated using AISA data, data from 15 *in situ* stations and data from three routine *in situ* stations, respectively.

Paper 3

The concentrations of chlorophyll-a (chl a), total suspended solids (TSS) and the absorption coefficient of colored dissolved organic matter (aCDOM(400)) are estimated in Case 2 waters using MERIS satellite (Level-1b) and AISA airborne spectrometer data acquired during a spring bloom in the Gulf of Finland, Baltic Sea. The accuracy of the estimation is analyzed using empirical band-ratio algorithms together with ground-based observations that include water samples analyzed in a laboratory (variation ranges: 22-130 μg/l, 2.9-20 mg/l, and 1.29-2.61 m^{-1} for chl a, TSS and aCDOM(400), respectively). Additional ground-based estimates (transects) on these characteristics are available through absorption and scattering coefficients measured with an ac-9 absorption and attenuation meter installed in a flow-though system. The retrieval accuracy (R^2) of all three water quality characteristics with MERIS data is close to or above 0.9, while the RMSE is 7.8 μg/l (22 %), 0.74 mg/l (16 %) and 0.08 m^{-1} (5 %), for chl a, TSS and aCDOM(400), respectively. The validity of the chl *a* algorithm is tested using nine additional data points. The BIAS-error for these points is 5.2 μg/l and the RMSE is 10.6 μg/l. The effects of changes in the atmospheric characteristics on band-ratio algorithms in cases where no concurrent in situ reference data are available are analyzed using the MODerate spectral resolution atmospheric TRANSmittance algorithm and computer model (MODTRAN). The additional error due to these changes is estimated to be below 20 % for the applied ratio algorithms. The water quality data available in the Level-2 MERIS-product distributed by the European Space Agency did not include valid results for the date investigated here.

Paper 4

This paper presents the results of a study where MODIS data and routinely collected *in situ* data were used to create a turbidity map of the lakes in southern Finland. The results also give an indication of the size of the buffer zone where the land surrounding the lakes disturbs the retrieval of water quality parameters. The retrieval of turbidity was based on empirical regression between *in situ* data and MODIS channel 1 (620-670 nm) radiance. The relationship is valid when turbidity ranges from 0 to 6 FNU. The accuracy of the retrieval reaches $R^2 = 0.9$ when the distance of pixel from shore is at least 450 m. The area where turbidity can be retrieved (after clouds, land areas and the 450-m buffer zone were masked) was about 10 % of the lake area covered by the MODIS image.

Paper 5

The use of airborne (AISA spectrometer) and simulated satellite (MERIS data simulated with AISA observations) remote sensing data for classification of three water quality variables: Secchi depth, turbidity and chlorophyll *a*, is studied. An extensive airborne spectrometer and ground truth data set obtained in four lake water quality measurement campaigns in southern Finland during 1996-1998 was used in the analysis. The class limits for the water quality variables were obtained from two operational classification standards. When remote-sensing data are used a combination of them proved to be the most suitable. The feasibility of the system for operational use was tested by training and testing the retrieval algorithms with separate data sets. In this case the classification accuracy is 90% for three Secchi

depth classes, 79% for five turbidity classes, and 78% for five chlorophyll *a* classes. When AISA data were spectrally averaged corresponding to Envisat-MERIS channels the classification accuracy was about the same as in the case of the original AISA channels.

Paper 6

The traditional method used in the water quality classification of Finnish lakes includes on site measurements and collection of water samples for subsequent analysis in laboratory conditions. The classification is based on statistical analysis of water quality parameter values and on expert opinion. It is possible to acquire similar information by using radiance values measured with the EOS Terra/Aqua MODerate resolution Imaging Spectroradiometer (MODIS). In this study, the classification accuracy with MODIS data is about 80%. Only about 0.2% of the 20,391 pixels were misclassified by two or more classes as a four-class classification system is used.

REFERENCES

Ahlman, M., Suomenlahden tila vuonna 1998 (In Finnish). Webpage: http://www.fimr.fi/fi/itamerikanta/bsds/253.html. Last updated: 3.9.2003. Visited: 24.10.2006.

Andersson, G.P., Kneizys, F.X., Chetwynd, J.H., Wang, J., Hoke, M.L., Rothman, L.S., Kimbali, L.M., McClathey, R.A., Shettle, E.P, Clough, S.A., Gallery, W.O., Abreu, L.W., and Selby, J.E.A., Fascode/Modtran/Lowtran: Past/Present/Future. Proceedings of 18th Annual Review Conference on Atmospheric Transmission Models, 6-8 June, 1995.

Antoine, D., and Morel, A., A multiple scattering algorithm for atmospheric correction of remotely sensed ocean colour (MERIS instrument): principle and implementation for atmospheres carrying various aerosols including absorbing ones. International Journal of Remote Sensing, vol. 20, no. 9, pp. 1875-1916, 1999.

Babey, S.K., and Anger, C.D., A compact airborne spectrographic imager (CASI). Proceedings of IEEE IGARSS'89, vol. 2, pp. 1028-1031, 1989.

Barnes, W.L., Pagano, T.S., and Salomonson, V.V., Prelaunch characteristics of the Moderate Resolution Imaging Spectroradiometer (MODIS) on EOS-AM1. IEEE Transactions on Geoscience and Remote Sensing, vol. 36, no. 4, pp. 1088-1100, 1998.

Bricaud, A., Morel, A. and Prieur, L., Absorption by dissolved organic matter of the sea (yellow substance) in the UV and visible domains. Limnology and Oceanography, vol. 26, no. 1, pp. 43-53, 1981.

Bricaud, A., Morel, A. and Prieur, L., Optical efficiency factors of some phytoplankters. Limnology and Oceanography, vol. 28, no. 5, pp. 816-832, 1983.

Buiteveld, H., Hakvoort, J.H.M. and Donze, M., The optical properties of pure water. Proceedings of the Ocean Optics XII Conference, SPIE, 2258, pp. 174–183, 1994.

Bukata, R.P., Jerome, J.H., Kondratyev, K.Y. and Pozdnyakov, D.V., Optical properties and remote sensing of inland and coastal waters, CRC Press, Boka Raton, Florida, USA, 362 p., 1995.

Bukata R.P., Satellite Monitoring of Inland and Coastal Water Quality. CRC press, Boka Raton, Florida, USA, 246 p., 2005.

Cocks, T., Jenssen, R., Stewart, A., Wilson, I., and Shields, T., The HyMap airborne hyperspectral sensor: the system, calibration and performance. 1st EARSeL Workshop on Imaging Spectroscopy, Zurich, Switzerland, EARSeL, Paris, pp. 37-42, 6-8 October 1998.

Cox, C. and W. Munk, Measurement of the roughness of the sea surface from photographs of the sun's glitter. Journal of the Optical Society of America, vol. 44, pp. 838-850, 1954.

D'Alimonte, D., Zibordi, G., Phytoplankton determination in an optically complex coastal region using a multilayer perceptron neural network. IEEE Transactions on Geoscience and Remote Sensing, vol. 41, pp. 2861-2868, 2003.

D'Sa, E. J., Miller, R. L., Bio-optical properties in waters influenced by the Mississippi River during low flow conditions. Remote Sensing of Environment, vol. 84, pp. 538-549, 2003.

Darecki, M., Weeks, A., Sagan, S., Kowalczuk, P., & Kaczmarek, S., Optical characteristics of two contrasting Case 2 waters and their influence on remote sensing algorithms. Continental Shelf Research, vol. 23, pp. 237–250, 2003.

Darecki, M., and Stramski, D., An evaluation of MODIS and SeaWiFS bio-optical algorithms in the Baltic Sea. Remote Sensing of Environment, vol. 89, pp. 326–350, 2004.

Dekker, A.G., Detection of optical water parameters for eutrophic lakes by high resolution remote sensing. PhD. Dissertation, Free University, Amsterdam, 1993.

Dekker, A.G., Malthus, T.J., and Seyhan, E. Quantitative modeling of inland water quality for high-resolution MSS systems. IEEE Transactions on Geoscience and Remote Sensing, vol. 29, no. 1, pp. 89 –95, 1991.

Dekker, A.G. and Peters, S.W.M., The use of the Thematic Mapper for the analysis of eutrophic lakes: A case study in The Netherlands. International Journal of Remote Sensing, vol. 14, no. 5, pp. 799-821, 1993.

Dekker, A.G., Vos, R.J., and Peters S.W.M., Comparison of remote sensing data, model results and *in situ* data for total suspended matter (TSM) in the southern Frisian lakes. The Science of the Total Environment, vol. 268, nos. 1-3, pp. 197-214, 2001.

EEA, European Environment Agency, European Rivers and Lakes: Summary, http://reports.eea.eu.int/87-90198-01-8/en/page002.html, 6.6.2006.

George, D.G. and Malthus, T.J., Using compact airborne spectrographic imager to monitor phytoplankton biomass in a series of lakes in north Wales. The Science of the Total Environment, vol. 268, nos. 1-3, pp. 215-226, 2001.

Giardino, C., Pepe, M., Brivio, P., Ghezzi, P. and Zilioli, E., Detecting chlorophyll, Secchi depth and surface temperature in sub-alpine lake using Landsat Imagery. The Science of the Total Environment, vol. 268, nos. 1-3, pp. 19-29, 2001.

Gitelson, A., The peak near 700 nm on radiance spectra of algae and water: relationships of its magnitude and position with chlorophyll concentration. International Journal of Remote Sensing, no. 13, pp. 3367-3373, 1992.

Gitelson, A., Garbuzov, G., Szilagyi, F., Mittenzwey, K-H., Karnieli, K., and Kaiser, A., Quantitative remote sensing methods for real-time monitoring of inland waters quality. International Journal of Remote Sensing, vol. 14, no. 7, pp. 1269-1295, 1993.

Gitelson, A., Yacobi, Y, Schalles, J, Ruinquist, D., Han, L., Stark, R., and Etzion, D., Remote estimation of phytoplankton density in productive waters. Archiv für Hydrobiologie, Advances in Limnology, vol. 55, pp. 121–136, 2000.

Gohin, F., Druon, J.N., Lampert, L., A five channel chlorophyll concentration algorithm applied to SeaWiFS data processed by SeaDAS in coastal waters. International Journal of Remote Sensing, vol. 23, pp. 1639-1661, 2002.

Gordon, H. R., Removal of atmospheric effects from satellite imagery of the oceans. Applied Optics vol. 17, no. 10, pp. 1631-1636, 1978.

Gordon, H.R., Brown, O.B., and Jacobs, M.M., Computed relationships between the inherent and apparent optical properties of a flat homogenous ocean. Applied Optics, vol. 14, pp. 417-427, 1975.

Gordon, H. R., and Clark, D. K., Remote sensing optical properties of a stratified ocean: An improved interpretation. Applied Optics, vol. 18, pp. 3428-3430, 1980.

Gordon, H., and A. Morel, Remote Assessment of Ocean Color for Interpretation of Satellite Visible Imagery: A Review. Lecture Notes on Coastal and Estuarine Studies, Vol. 4, Springer Verlag, New York, 114 pp, 1983.

Gregg, W. W., Casey, N. W., Global and regional evaluation of SeaWiFS chlorophyll data set. Remote Sensing of Environment, vol. 93, pp. 463-479, 2004

Hallikainen, M., Kemppinen, M., Philflyckt, J., Mononen, I., Auer, T., Rautiainen, K., Lahtinen, J., Tauriainen, S., Valmu, H., HUTRAD: Airborne multifrequency microwave radiometer. Proceedings of 2nd ESA Workshop on Millimetre Wave Technology and Applications: Antennas, Circuits and Systems, pp. 115-120, Espoo, Finland, 27-29 May 1998.

He, M-X., Liu, Z-S., Du, K-P., Li, L-P., Chen, R., Carder, K.L., Lee, Z-P., Retrieval of chlorophyll from remote-sensing reflectance in the China seas. Applied Optics, vol. 39, no 15, pp. 2467-2474, 2000.

Heinonen, P., and Herve, S., Water quality classification of inland waters in Finland. Aqua Fennica, vol. 17, no. 2, pp. 147-156, 1987.

Henriksen, A., Skjelvåle, B.L., Mannio, J., Wilander, A., Jensen, J.P., Moiseenko, T., Harriman, R., Traaen, T.S., Fjeld, E., Vuorenmaa, J., Kortelainen, P. & Forsius, M., Results of National Surveys (1995) in Finland, Sweden, Russian Kola, Russian Karelia, Scotland and Wales. Norwegian Institute for Waters Research NIVA, Report SNO 3645-97, 43 pp., 1997.

Herlevi, H., Inherent and apparent optical properties in relation to water quality in Nordic waters. PhD Thesis, University of Helsinki, Report Series in Geophysics, 45, 2002.

Hovis, W.A., Clark, D.K., Anderson, F., Austin, R.W., Wilson, W.H., Baker, E.T., Ball, D., Gordon, H.R., Mueller, J.L., EI-Sayed, S.Z., Strum, B., Wrigley, R.C., and Yentsch, C.S., Nimbus-7 Coastal Zone Color Scanner: System description and initial imagery. Science, vol. 210, pp. 60-63, 1980.

Hu, C., Chen, Z., Clayton, T.D., Swarzenski, P., Brock, J.C., Muller-Karger, F.E., Assessment of estuarine water-quality indicators using MODIS medium-resolution bands: Initial results from Tampa Bay, FL. Remote Sensing of Environment, vol. 93, pp. 423-441, 2004.

Härmä, P., Vepsäläinen, J., Hannonen, T., Pyhälahti, T., Kämäri, J., Kallio, K., Eloheimo, K., and Koponen, S., Detection of water quality using simulated satellite data and semi-empirical algorithms in Finland. The Science of the Total Environment, vol. 268, nos. 1-3, pp. 107-122, 2001.

Iluz, D., Yacobi, Y.Z., Gitelson, A., Adeptation of an algorithm for chlorophyll-a estimation by optical data in the oligotrophic Gulf of Eilat. International Journal of Remote Sensing, vol. 24, pp. 1157-1163, 2003.

ISO 7027. Water quality. Determination of turbidity. International Standards Organization, 1990.

ISO 10260. Water quality – Measurement of biochemical parameters – Spectrometric determination of the chlorophyll a concentration. International Organization for Standardization, 1992.

Jerlov, N.G., Marine Optics. Elsevier, Amsterdam, 1976.

Jorgensen P.V., Standard CZCS Case 1 algorithms in Danish coastal waters. International Journal of Remote Sensing, vol. 20, no. 7, pp. 1289-1301, 1999.

Jorgensen P.V., SeaWiFS data analysis and match-ups with in situ chlorophyll concentrations in Danish waters. International Journal of Remote Sensing, vol. 25, no. 7–8, pp. 1397–1402, 2004.

Kallio, K., Kutser, T., Hannonen, T., Kutser, T., Koponen, S., Pulliainen, J., Vepsäläinen, J., and Pyhälahti, T. Retrieval of water quality variables from airborne spectrometer in various lake types at different seasons. The Science of the Total Environment, vol. 268, nos. 1-3, pp. 59-78, 2001.

Kallio, K., Pulliainen, J. and Ylöstalo, P., MERIS, MODIS and ETM+ Channel Configurations in the Estimation of Lake Water Quality from Subsurface Reflectance Using Semianalytical and Empirical Algorithms. Geophysica, vol. 41, nos. 1-2, pp. 31-55, 2005.

Kallio, K., Optical properties of Finnish lakes estimated with simple bio-optical models and water quality monitoring data. Nordic Hydrology vol. 37, pp. 183-204, 2006.

Kirk, J.T.O., Light and photosynthesis in aquatic ecosystems. Cambridge University Press, Melbourne, Australia, 401 p, 1983.

Kirk, J.T.O., Light and photosynthesis in aquatic ecosystems. Cambridge University Press, U.K., 509 p, 1994.

Kirk, J.T.O., Dependence of relationship between inherent and apparent optical properties of water on solar altitude. Limnology and Oceanography, vol. 29, no. 2, pp. 350-356, 1984.

Kutser, T., Quantitative detection of chlorophyll in cyanobacterial blooms by satellite remote sensing. Limnology and Oceanography, vol. 49, no. 6, pp. 2179–2189, 2004.

Kristensen, P., and Hansen, H.O. (editors), European rivers and lakes; Assessment of their environmental state. European Environmental Agency, Copenhagen, 122 p., 1994.

Lindell, T., Pierson, D., Premazzi, G., and Zilioli, E. (editors), Manual for Monitoring European Lakes Using Remote Sensing Techniques. Luxembourg: Office for Official Publications of the European Communities, 164 p., 1999.

Lindfors, A., Rasmus K., and Strömbeck N., Point or pointless – quality of ground data. International Journal of Remote Sensing, vol. 26(2), pp. 415-423, 2005.

Maffione, R.A. and Dana, D.R., Recent measurements of the spectral backward-scattering coefficient in coastal waters. Proceedings of the Ocean Optics XIII, SPIE, 2963, pp. 154–159, 1996.

Mayo, M., Gitelson, A., Yacobi, Y.Z., and Ben-Avraham, Z., Chlorophyll distribution in Lake Kinneret determined from Landsat Thematic Mapper data. International Journal of Remote Sensing, vol. 16, no. 1, pp. 175-182, 1995.

Mélin, F., Zibordi, G., Berthon, J-F., Assessment of SeaWiFS Atmospheric and Marine Products for the Northern Adriatic Sea. IEEE Transactions on geoscience and remote sensing, vol. 41, pp. 548-558, 2003.

Measures, R.M., Laser Remote Sensing, fundamentals and applications. New York, John Wiley & Sons. 510 p., 1984.

Miller, R.L., and McKee, B.A., Using MODIS Terra 250 m imagery to map concentrations of total suspended matter in coastal waters. Remote Sensing of Environment, vol. 93, pp. 259-266, 2004.

Milton, J.S. and Arnold, J.C., Introduction to probability and statistics: Principles and applications for engineering and computing sciences. McGraw-Hill Publishing Company, Singapore, 811 p., 1995.

Mobley, C.D., Light and water; Radiative transfer in natural waters. Academic Press, Inc., San Diego, 592 p., 1994.

Mobley, C.D., Stramski, D., Bissett, W.P., and Boss, E., Optical modeling of ocean waters: Is the Case 1–Case 2 classification still useful? Oceanography, vol. 17, no. 2, pp. 60–67, 2004.

MODIS, Internet webpage, http://modis.gsfc.nasa.gov/, 6.4.2006.

Moore, G., Aiken, J., and Lavender, S., The atmospheric correction of water colour and the quantitative retrieval of suspended particulate matter in Case II waters: application to MERIS. International Journal of Remote Sensing, vol. 20, no. 9, pp. 1713-1733, 1999.

Moore, J.K., Abbott, M.R., Richman, J.G., Smith, W.O., Cowles. T.J., Coale, K.H., Gardner, W.D., and Barber, R.T., SeaWiFS satellite ocean color data from the Southern Ocean. Geophysical Research Letters, vol. 26, no. 10, pp. 1465-1468, 1999.

Morel, A. and Prieur, L., Analysis of variations in ocean color. Limnology and Oceanography, vol. 22, pp. 709-722, 1977.

Morel, A., Optical modeling of the upper ocean in relation to its biogeneous matter content (Case 1 waters). Journal Geophysical Research., vol. 93(C9), pp. 10749-10768, 1988.

Murphy, J. and Riley, J. P., A modified single solution method for the determination of phosphate in natural waters. Analytica Chimica Acta, vol. 27, pp. 31-36, 1962.

Mäkisara, K., Meinander, M., Rantasuo, M., Okkonen, J., Aikio, M., Sipola, K., Pylkkö, P., and Braam, B., Airborne Imaging Spectrometer for Applications (AISA). Proceedings of IEEE IGARSS'93, 18-21 August, Tokyo, Japan, pp. 479-481, 1993.

NASA, Ocean Color From Space - Measurement Technique. Internet webpage: http://disc.gsfc.nasa.gov/oceancolor/scifocus/space/ocdst_measurement_technique.shtml., Last updated: 14.2.2004. Visited: 27.10.2006.

Nechad, B., De Cauwer, V., Park, Y., Ruddick, K.G., Suspended particulate matter (SPM) mapping from MERIS imagery. Calibration of a regional algorithm for the Belgian coastal waters. Proceedings of MERIS User Workshop, Frascati, Italy, 10-13 November 2003.

OECD, Eutrophication of water, monitoring, assessment and control. Organization for Economic Cooperation and Development. Paris, 150 p., 1982.

Perttilä, M., Hällfors, M., and Koivula, M., The shape, area and volume of the Baltic Sea. Internet webpage: http://www.fimr.fi/en/itamerikanta/bsds/1709.html. Last updated: 4.10.2004. Visited: 24.10.2006.

Pope, R.M. and Fry, E.S., Absorption spectrum (380 – 700 nm) of pure water: II. Integrating cavity measurements. Applied Optics, vol. 36, pp. 8710-8723, 1997.

Premazzi, G. and Chiaudani, G., Ecological quality of surface waters, Quality assessment schemes for European community lakes. Brussels – Luxembourg, Commission of the European Communities. 124 p., 1992.

Pulliainen, J., Vepsäläinen, J., Kallio, K., Koponen, S., Pyhälahti, T., Härmä, P., Hallikainen, M. Monitoring of water quality in lake and coastal regions using simulated ENVISAT MERIS data. Proceedings of ESA ERS-ENVISAT Symposium, pp. 10, Göteborg, 16-20 October 2000.

Pulliainen, J., Kallio, K., Eloheimo, K., Koponen, S., Servomaa, H., Hannonen, T., Tauriainen, S., and Hallikainen, M., A semi-operative approach to water quality retrieval from remote sensing data. The Science of the Total Environment, vol. 268, nos. 1-3, pp. 79-94, 2001.

Raatikainen, M. and Kuusisto, E., The number and surface area of lakes in Finland. In Finnish, with English Abstract, Terra, vol. 102, pp. 97-110, 1988.

Rast, M., Bézy, J., and Bruzzi, S., The ESA Medium Resolution Imaging Spectrometer MERIS – a review of the instrument and its mission. International Journal of Remote Sensing, vol. 20, no. 9, pp. 1681-1702, 1999.

Roesler, C.S., Perry, M.J. and Carder, K.L., Modeling in situ phytoplankton absorption from total absorption spectra. Limnology and Oceanography, 34, 1512–1525, 1989.

Sathyendranath, S., Prieur, L. and Morel, A., A three component model of ocean colour and its application to remote sensing of phytoplankton pigments in coastal waters. International Journal of Remote Sensing, 10, 1373–1394, 1989.

Schalles, J.F., Gitelson A.A., Yacobi Y.Z., and Kroenke A.E., Estimation of chlorophyll *a* from time series measurements of high spectral resolution reflectance in a eutrophic lake. Journal of Phycology, vol. 34, pp. 383-390, 1998.

Shifrin, K.S. Physical Optics of Ocean Water, American Institute of Physics, New York, 285 p., 1988.

Smith, R.C., and Baker, K.S., Optical classification of natural waters. Limnology and Oceanography, vol. 23, pp. 260-267, 1978.

Strömbeck, N., Candiani, G., Giardino, C., Zilioli, E., Water quality monitoring of lake Garda using multi-temporal MERIS data. Proceedings of MERIS User Workshop, Frascati, Italy, 10-13 November 2003.

Swedish Environmental Protection Agency (EPA), http://www.internat.environ.se/documents/legal/assess/assedoc/lakes.htm, 6.6.2006.

SYKE, Finnish Environment Institute, Water quality monitoring Internet homepage. http://www.ymparisto.fi/default.asp?contentid=152352&lan=fi&clan=en, 6.6.2006.

Särkkä, J., Järvet ja Ympäristö: limnologian perusteet (In Finnish). Helsinki, Gaudeamus, 157 p., 1996.

Tang, D., Kawamura, H., Lee, M-A., Van Dien, T., Seasonal and spatial distribution of chlorophyll-a concentrations and water conditions in the Gulf of Tonkin, South China Sea. Remote Sensing of Environment, vol. 85, pp. 475-483, 2003.

Thiemann, S., and Kaufmann, H., Determination of chlorophyll content and trophic state of lakes using field spectrometer and IRS-1C satellite data in the Mecklenburg

lake district, Germany. Remote Sensing of Environment, vol. 73, no. 2, pp. 227-235, 2000a.

Thiemann, S., and Kaufmann, H., Lake water quality monitoring using hyperspectral airborne data - a multitemporal approach. Proceedings of the Sixth International Conference on Remote Sensing for Marine and Coastal Environments, Charleston, South Carolina, USA, vol. II, pp. 157-164, 1-3 May 2000b.

Thiemann, S., personal communication, 15 October 2001.

Thiemann, S., Kaufmann, H., Lake water quality monitoring using hyperspectral airborne data—a semiempirical multisensor and multitemporal approach for the Mecklenburg Lake District, Germany. Remote Sensing of Environment, vol. 81, pp. 228-237, 2002

Ulloa, O., Sathyendranath, S., Platt, T., and Quiñones, R.A., Light scattering by marine heterotrophic bacteria. Journal of Geophysical Research, vol. 97, pp. 9619-9629, 1992.

Vepsäläinen J., Pulliainen J., Kallio K., Pyhälahti T., Koponen S., Lindfors A., Rasmus K., Atmospheric Correction for Case 2 Waters with a Bio-Optical Reflectance Model, Proceedings of 31st International Symposium on Remote Sensing of Environment, Saint Petersburg, Russia, June 2005.

Vermote, E.F., Tanre, D., Deize, J.L., Herman, M., and Morcrette, J.J., Second simulation of the satellite signal in the solar spectrum, 6S: an overview. IEEE Transaction on Geoscience and Remote Sensing, vol. 35, pp. 675-686, 1997.

Vincent, R.K., Qin, X., McKay, R.M.L., Miner, J., Czajkowski, K., Savino, J., Bridgeman, T., Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie. Remote Sensing of Environment, vol. 89, pp. 381-392, 2004

Vuoristo, H., Water quality classification of Finnish inland waters. European Water Management, vol. 1, no.6, pp. 35-41, 1998.

Woodruff, D.L., Stumpf, R.P., Scope, J.A., and Paerl, H.W., Remote estimation of water clarity in optically complex estuarine waters. Remote Sensing of Environment, vol. 68, no. 1, pp. 41-52, 1999.

Yentsch, C.S. and Yentsch, C.M., Fluorescence spectral signatures: the characterization of phytoplankton populations by the use of excitation and emission spectra. Journal of Marine Research, vol. 37, no. 3, pp. 471-483, 1979.

Zibordi, G., Mélin, F., Berthon, J-F., van der Linde, D., Bulgarelli, B., An evaluation of the accuracy of MERIS radiometer and geophysical products for a northern adriatic sea coastal site. Proceedings of MERIS User Workshop, Frascati, Italy, 10-13 November 2003.

Zoloratev, V.M., and Demin, A.V., Optical constants of water over a broad range of wavelengths, 0.1 Å – 1 m, Optical spectroscopy USSR, vol. 43, no. 2, p 157, 1977.

Ylöstalo, P., Kallio, K. and Seppälä, J., Absorption properties of particles and CDOM in lake waters, (in preparation), 2006.

Ympäristö, Lakes in Finland. Internet webpage: http://www.ymparisto.fi/default.asp?contentid=171561&lan=fi&clan=en. Last updated: 21.2.2006. Visited: 24.10.2006.

Ympäristö, Kosteikot pohjoismaissa ja Ramsar-sopimus (In Finnish) 12 p., 2004. Also available from: http://www.ymparisto.fi/download.asp?contentid=35753&lan=fi

Ympäristöministeriö, Suomen rannikkostrategia (in Finnish), Suomen ympäristö 10/2006, Luonnonvarat, 82 p., 2006. Also available from: http://www.ymparisto.fi/download.asp?contentid=52277&lan=fi.

Östlund, C., Flink, P., Strömbeck, N., Pierson, D., and Lindell, T., Mapping of the water quality of Lake Erken, Sweden, from imaging spectrometry and Landsat Thematic Mapper. The Science of the Total Environment, vol. 268, nos. 1-3, pp. 139-154, 2001.

APPENDIX A

This appendix contains a brief description of the bio-optical reflectance model used in [P6]. For a more complete description see Kallio et al. (2005) and Kallio (2006).

Four optically active components are assumed in the model: phytoplankton, tripton, coloured dissolved organic matter and pure water. The total spectral absorption coefficient ($a_{Tot}(\lambda)$) is described by:

$$a_{Tot}(\lambda) = a_w(\lambda) + a_{cdom}(\lambda) + a_{Ph}(\lambda) + a_{Tri}(\lambda), \quad (1)$$

where $a_w(\lambda)$ is the absorption coefficient of pure water (Buiteveld 1994), $a_{ph}(\lambda)$ is the absorption coefficient of phytoplankton, $a_{cdom}(\lambda)$ is the absorption coefficient of CDOM and $a_{Tri}(\lambda)$ is the specific absorption coefficient of tripton.

Absorption by CDOM is calculated by assuming an exponential increase with a decreasing wavelength (Bricaud *et al.* 1981):

$$a_{cdom}(\lambda) = a_{cdom}(400)\, e^{-S_{cdom}(\lambda-400)}, \quad (2)$$

where $a_{cdom}(400)$ is the absorption coefficient of CDOM at 400 nm and S_{cdom} is the slope factor.

Absorption by phytoplankton, $a_{ph}(\lambda)$, is calculated by:

$$a_{Ph}(\lambda) = a^*_{ph}(\lambda) C_{Chl-a}, \quad (3)$$

where $a^*_{ph}(\lambda)$ is the Chl-a specific absorption coefficient of phytoplankton. $a^*_{ph}(\lambda)$, decreases in the model with increasing $C_{Chl\text{-}a}$ according to measurements made in Finnish lakes (Ylöstalo *et al.* 2005).

Absorption by tripton, $a_{Tri}(\lambda)$, is expressed as:

$$a_{Tri}(\lambda) = a^*_{\underline{TSS}}(\lambda) C_{TSS}, \quad (4)$$

where $a^*_{\underline{TSS}}(\lambda)$ is the specific absorption of bleached total suspended solids and C_{TSS} is the concentration of TSS. Absorption by tripton is defined using C_{TSS}, because tripton concentration measurements were not available. $a^*_{\underline{TSS}}(\lambda)$ is described analogously to the calculation of $a_{cdom}(\lambda)$ (Roesler *et al.* 1989):

$$a^*_{\underline{TSS}}(\lambda) = a^*_{\underline{TSS}}(400) e^{-S_{Tri}(\lambda-400)}, \quad (5)$$

where $a^*_{\underline{TSS}}(400)$ is the specific absorption of bleached total suspended solids at 400 nm and S_{Tri} is the slope factor of tripton absorption.

The total scattering coefficient, $b_{tot}(\lambda)$, is described by:

$$b_{tot}(\lambda) = b_w(\lambda) + b^*_{TSS}(\lambda) C_{TSS}, \quad (6)$$

where b_w is the scattering coefficient of pure water (Buiteveld 1994), and b^*_{TSS} is the specific scattering coefficient of TSS.

The specific scattering coefficient of TSS ($b^*_{TSS}(\lambda)$) is described by a power function (*e.g.* Maffione and Dana 1996, Herlevi 2002):

$$b^*_{TSS}(\lambda) = b^*_{TSS}(555) \left(\frac{555}{\lambda}\right)^{n_b}, \quad (7)$$

where $b^*_{TSS}(555)$ is the specific scattering coefficient of TSS at 555 nm and n_b is the scattering exponent.

The total attenuation coefficient is the sum of the absorption and scattering coefficients:

$$c_{Tot}(\lambda) = a_{Tot}(\lambda) + b_{Tot}(\lambda), \quad (8)$$

The total backscattering coefficient, $b_{b,tot}(\lambda)$, needed in the reflectance model, is described by:

$$b_{b,tot}(\lambda) = bp_w b_w(\lambda) + bp_{TSS} b^*_{TSS}(\lambda) C_{TSS}, \quad (9)$$

where b_w is the scattering coefficient of pure water and b^*_{TSS} is the specific scattering coefficient of TSS. bp_w and bp_{TSS} are the backscattering ratios of pure water and TSS, respectively. The coefficients of the bio-optical model are summarized in Table A1.

Reflectance model

Calculation of irradiance reflectance just beneath the water surface, R(0-,λ), is based on the following equation (Gordon *et al.* 1975 and simplified by Jerlov 1976):

$$R(0-,\lambda) = C\frac{b_{b,Tot}(\lambda)}{a_{Tot}(\lambda) + b_{b,Tot}(\lambda)}, \quad (10)$$

where $b_{b,Tot}(\lambda)$ is the total backscattering coefficient and $a_{Tot}(\lambda)$ is the total absorption coefficient. The coefficient C depends mainly on the illumination and the viewing geometry. C is estimated by the equation presented by Kirk (1984):

$$C = -0.629\mu_0 + 0.975, \quad (11)$$

where μ_0 is the cosine of the solar zenith angle in the water. The solar zenith angle used in all R(λ) calculations was 46.3° (1st of August, 13.00 local time, latitude 64° N). $a_{Tot}(\lambda)$ and $b_{b,Tot}(\lambda)$ are obtained by summing up the absorption and backscattering coefficients of the optically active substances in the water (Eqs. (1) and (9)).

Table A1. Coefficients of the bio-optical reflectance model.

Coefficient	Symbol	Value	Source
Absorption coefficient of pure water	$a_w(\lambda)$	see the reference	Buiteveld (1994)
Specific absorption of phytoplankton	$a^*_{ph}(\lambda)$	see the reference	Ylöstalo *et al.* (2005)
Slope factor of CDOM absorption	S_{CDOM}	0.0150 nm^{-1}	Kallio *et al.*(2005)
Specific absorption of bleached TSS at 400 nm	$a^*_{TSS}(400)$	0.13 l $m^{-1}mg^{-1}$	Ylöstalo *et al.* (2005)
Slope factor of tripton absorption	S_{tri}	0.012 nm^{-1}	Ylöstalo *et al.* (2005)
Specific scattering of TSS at 555 nm	$b^*_{TSS}(555)$	0.811 l $m^{-1}mg^{-1}$	Kallio *et al.(*2005)
Scattering exponent of TSS	n_b	0.705	Kallio *et al.* (2005)
Scattering coefficient of pure water	$b_w(\lambda)$	see the reference	Buiteveld (1994)
Backscattering probability of pure water	bp_w	0.5	Sathyendranath *et al.* (1989)
Backscattering probability of TSS	bp_{TSS}	0.0131	Kallio *et al.* (2005)